Micro Shocks and Macro Fluctuations in the Information Network

Zhen Huo
Yale University

Minghao Li
Peking University

Jieran Wu
Zhejiang University

Huan Xiong
Harbin Institute of Technology

March 20, 2025

Abstract

Traditional granular macroeconomic models rely on heterogeneities in firm sizes (Gabaix, 2011) or importance in input-output linkages Acemoglu et al. (2012). In this paper, we explore the role of another source of heterogeneity in amplifying the impact of micro shocks on aggregate fluctuations—firms' attention allocation. Empirically, we document that the attention firms receive follows a fat-tailed distribution based on their browsing activities related to each other's electronic filings on EDGAR. This suggests that a small number of firms receive disproportionately large amounts of attention. The received attention is positively correlated with firm sizes, but not perfectly. Theoretically, we extend the noisy business cycle framework (Angeletos and La'O, 2010) to incorporate asymmetric attention at the firm level through a directed graph. We provide the condition on the in-degree and out-degree of this information network that allows for granular effects, even with homogeneous firm sizes. A single firm can have an aggregate impact only if a diluted in-degree index is sufficiently large. The mechanism is that, with incomplete information, salient firms can facilitate coordination as semi-public signals. Quantitatively, we show that the attention channel alone can generate sizable aggregate fluctuations driven by micro shocks. However, whether the asymmetric attention channel complements the heterogeneous firm size channel can be ambiguous and hinges on the correlation between firm size and received attention.

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

1. Introduction

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

This paper makes three contributions in addressing these questions: Empirically, we document that the attention firms receive follows a fat-tailed distribution based on their browsing activities related to each other's electronic filings on EDGAR (Electronic Data Gathering, Analysis, and Retrieval). This suggests that a small number of firms receive disproportionately large amounts of attention. In addition, the most salient firms may not necessarily be the largest firms. Theoretically, we extend the noisy business cycle framework (Angeletos and La'O, 2010) to incorporate asymmetric attention at the firm level through a directed graph. We provide the condition on the in-degree and out-degree of this information network that allows for granular effects, even with homogeneous firm sizes. Quantitatively, we show that the attention channel alone can generate sizable aggregate fluctuations driven by micro shocks. However, whether the asymmetric attention channel complements the heterogeneous firm size channel can be ambiguous and hinges on the strength of the general equilibrium considerations.

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

We document three facts. First, the distributions of browsing received and sent both exhibit power laws. The browsings received follow almost an exact Zipf's law. The estimated tail parameter for the browsings sent is -0.78. That is, a small number of firms receive a disproportionately large amount of attention. Second, the browsing activities are weakly positively correlated with firm sizes. One may have the prior that the firms that receive most browsing requests are those biggest firms. However, the correlation between the browsings received and the sales share is only 0.29. Fewer than 20 firms that belong to the top 100 largest firms also belong to the top 100 firms that receive the most browsing. Third, the browsing received and sent by a firm are positively correlated only for large firms. Outside

the top 400 largest firms, the browsing received and sent become uncorrelated.

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

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

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

Under what conditions can a firm's firm-specific component induce aggregate fluctuations when the number of firms goes to infinity? We answer this question in two steps. First, we define a key statistics in gauging firms' importance in the information network, a diluted in-degree index. For an firm i, the index is given by the sum of reciprocal of other firms' out-degrees which observes firm i. This index summarizes the effective attention received. Second, we provide an estimate of the bounds of the GDP elasticity with respect to a firm's micro shock, both of which are hinges on the diluted in-degree index. The granular effects can emerge if only if such index is growing with the number of firms at the same rate.

Intuitively, the granular effects can happen if a firm i is salient in the economy in the sense that it receives attention from most of other firms in the economy. However, this is not sufficient. It has to be that such attention is concentrated on firm i and few other firms can share this attention. For example, granularity cannot take place when a fixed fraction of total firms are observed by everyone. In contrast, when the in-degrees follow a Zipf's law distribution, granularity will emerge.

Quantification and Implications. To quantify the information channel in amplifying the effects of micro shocks, we adopt a random graph appraoch when calibrating the information matrix. The leverage on the patterns of the browsing activities in EDGAR to assign probabilities for an element in the information matrix to be 1. When a firm i engages in more browsing activities and when a firm j receives more browsing requests, the aforemention probability becomes higher. We further extends

our theoretical framework to allow heterogeneity in firm sizes and the volatility of the firm-specific compnenets.

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

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

We also provide two testable implications. First, in our model, salient firms contribute more to aggregate fluctuations, and therefore are more correlated with GDP. We show that in the data, the correlation between firms' quarterly sales and GDP is higher for firms that receive more browsing requests, conditional on the average firm sizes. Second, in our model salient firms have more influence on other firms' beliefs. As a result, firms' expectations about their own sales' growth should be more correlated with the expectations of the salient firms.

Related literature. The granular effect in macroeconomics, introduced by Gabaix (2011), posits that idiosyncratic shocks to large firms may have significant macroeconomic implications, challenging the conventional assumption that firm-level shocks average out in the aggregate. There are two strands of literature that provide micro-foundation for this granular effect. The first strand focuses on the heterogeneity in the distribution of firm size and market power. Gabaix's seminar work demonstrates that in economies with fat-tailed firm-size distributions, firm-specific shocks propagate to the macroeconomy and induce aggregate fluctuations. Carvalho and Gabaix (2013) extend the "granular hypothesis" by showing that the evolving patterns of sectoral concentrations amplify the impact of granular shocks and account for the dynamics of aggregate volatility. More recently, Eeckhout et al. (2020) highlights how the growing dominance of large firms affects productivity and business cycle dynamics. The second strand of literature explores the role of economy's input-output

linkages, or production network, in translating idiosyncratic shocks into sizable macroeconomic fluctuations. Acemoglu et al. (2012), Acemoglu et al. (2016), and Baqaee and Farhi (2020) adopt this approach by emphasizing the role of intersectoral dependencies in propagating firm-level shocks through supply chains, thereby amplifying their macroeconomic effects or creating cascading failures (Baqaee, 2018). Relative to the literature, we propose a novel channel of granularity effect that relies on the information network among firms in the economy. Supported by the empirical evidence, we show that firms' idiosyncratic shocks are able to generate sizable aggregate fluctuations that are independent of size-heterogeneity and production-network structure.

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

2. Motivating Facts

This section documents three novel facts about firms' browsing activities using internet logs of firms' access to others' filings on EDGAR. These facts motivate our subsequent theoretical analysis:

1. The distributions of browsing received and sent both exhibit power-law behavior.
2. Browsing activities are weakly positively correlated with firm size.
3. Browsing received and sent are positively correlated only for large firms.

Before presenting these facts in detail, we first describe the data.

2.1 The EDGAR Browsing Data

Our motivating facts exploit the detailed firm-to-firm browsing information on documents filed at EDGAR, which stands for the Electonic Data Gathering, Analysis, and Retrieval system. The EDGAR

system is an electronic platform maintained by the U.S. Securities and Exchange Commission (SEC).[1] Since 1996, all U.S. public companies are required to file disclosures electronically on EDGAR. These filings, freely accessible online, are a primary source to obtain other firms' information for managers, investors, and analysts.[2] Table 6 in the Appendix lists the most browsed document types.

Starting in 2003, the SEC's Division of Economic and Risk Analysis (DERA) began publishing internet traffic data for SEC filings, known as the "EDGAR Logfile." This dataset includes detailed information on the online viewing record for every SEC filing published on the EDGAR system, which includes a partially masked IP addresses, access timestamps, and unique document identifiers (Accession numbers).

Measuring browsing activity. We define a *browser* as the viewer of a document and a *browsee* as the document owner. Using the method from Chen et al. (2020), we unmask the IP addresses and map them to firm identities via ip-info.io, a leading IP information provider. This yields a daily browser-browsee database.

We define browsing intensity, denoted as b_{ijt}, to measure the number of accesses by browser i to browsee j's filings during time interval t:

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

where $a_{ijs} = 1$ if browser i accesses browsee j's filings at time s. We define firm i's total browsing received during time interval t as $B^R_{it} = \sum_j b_{jit}$, and define firm i's total browsing sent during time interval t as $B^S_{it} = \sum_j b_{ijt}$.

Baseline Sample. Our baseline sample focuses on public firms, as they are the primary filers on EDGAR and their data can be matched with sources like Compustat. We restrict the sample period from 2009 to 2016.[3] The sample includes 713,157,510 unique IP addresses from 7,622 public companies, representing 52.4% of all EDGAR browsings during the sample period.[4]

On an annual average basis, there are 3,626 unique browsers and 6,254 unique browsees; a browser views 180 browsees and a browsee is viewed by 105 browsers. The annual browsing volume is 60 for a browser-browsee pair. Finally, the browsing volume of a typical browser amounts to 11,356, and the browsing volume received by a typical browsee is 5,897.

Table 1 summarizes our baseline sample. Annually, there are 3,626 unique browsers and 6,254 unique browsees, with a browser on average viewing 180 browsees and a browsee viewed by 105 browsers. The average annual browsing volume is 60 per browser-browsee pair. Finally, the total

[1] The SEC website provides more information on the EDGAR system: https://www.sec.gov/edgar/about.

[2] EDGAR filings provide timely and comprehensive information, such as Form 8-K for unscheduled events and Form 10-K for annual updates. EDGAR is also the primary source for filings, as it offers free access and retains critical information often lost on other platforms like Bloomberg or Yahoo Finance.

[3] We start in 2009 due to limited pre-2009 data and end in 2017 when SEC stopped releasing traffic data.

[4] We exclude automated crawlers using the method in Cao et al. (2023). Our results are robust to their inclusion.

browsing volume of a typical browser amounts to 11,356, and the total browsing volume received by a typical browsee is 5,897.

Table 1: EDGAR Browsing Data: Summary Statistics

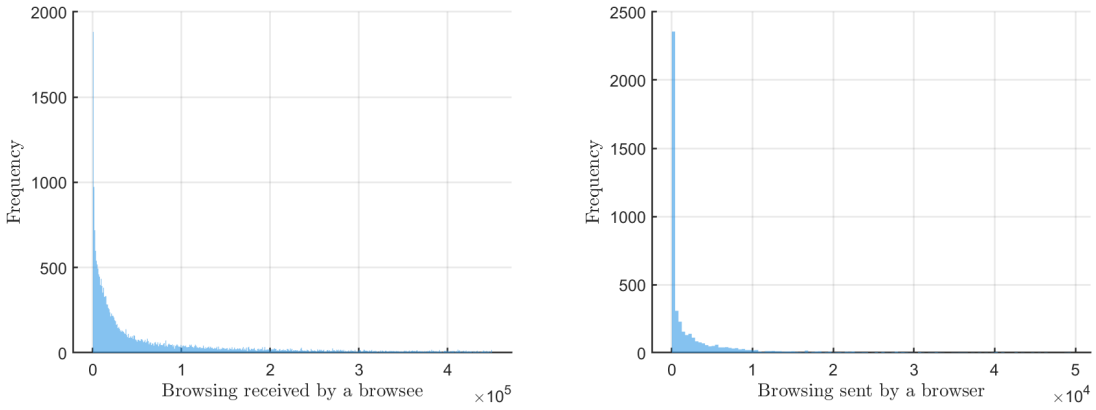
	Mean	Median	S.D.
No. Browsers per Year	3,626	3,608	251
No. Browsees per Year	6,254	6,280	321
No. Browsees viewd by a Browser per Year	180	176	25
No. Browsers viewing a Browsee per Year	105	74	115
Browsing Volume of a Browser per Year	11,356	9,881	7,764
Browsing Volume received by a Browsee per Year	5,897	5,389	4,869
Browsing Volume per Browser-Browsee pair per Year	60	50	34

2.2 Three Facts on Firms' Browsing Behavior

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

To begin with, Figure 1 plots the distributions of browsing received and sent in 2016. Both distributions are highly right-skewed, indicating that while most firms receive or send a small number of browsings, a small fraction of firms account for a disproportionately large share of activity. The long tail to the right of the distributions suggests that both distributions are fat-tailed, which possibly follow the power law.

Figure 1: Distribution of Firms' Browsings



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

To gain a further understanding of the distributions, we first rank firms based on their browsings received or sent in 2016.[5] We then obtain the Rank-1/2 estimator proposed in Gabaix and Ibragimov (2011) by regressing the log rank of firms on their log browsing activity,[6]

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

where Rank_i is the rank of firm i and B_i denotes firm i's total browsing received or sent. The coefficient α represents the power-law exponent.

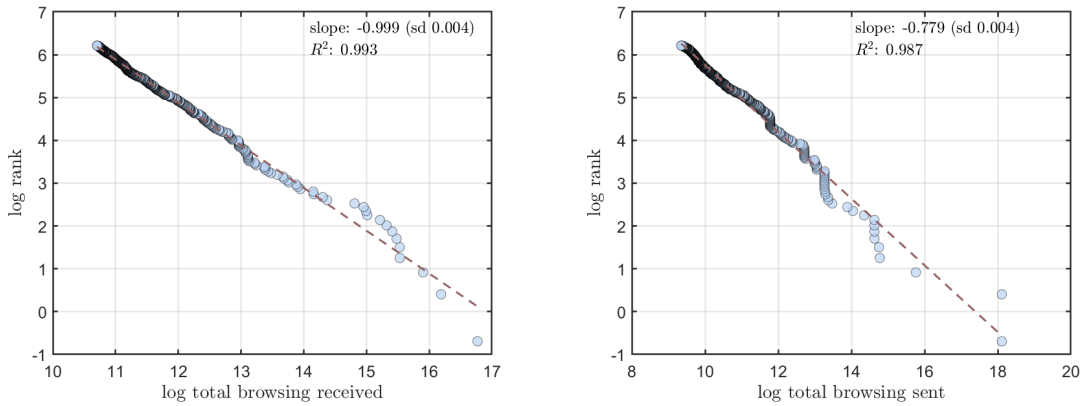
For browsing received, the estimated α is equal to -0.999. The standard deviation is 0.004, and the R^2 is 0.993. This estimate indicates that the distribution of firms' browsings received follow Zipf's law. That is, the probability that the browsing received exceeds a threshold K is proportional to $1/K$,

$$\text{Prob}(B_i > K) = a/K^{\alpha}, \text{ with } \alpha \approx 1,$$

where a is a positive constant. This Zipf's law is visualized in the left panel of Figure 5.

For browsing sent, the estimated α is equal to -0.779 with standard deviation 0.004 and R^2 equal to 0.987, suggesting a fatter tail than for browsing received. The right panel of Figure 5 illustrates this power-law behavior.

Figure 2: The Power Law



Note:

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

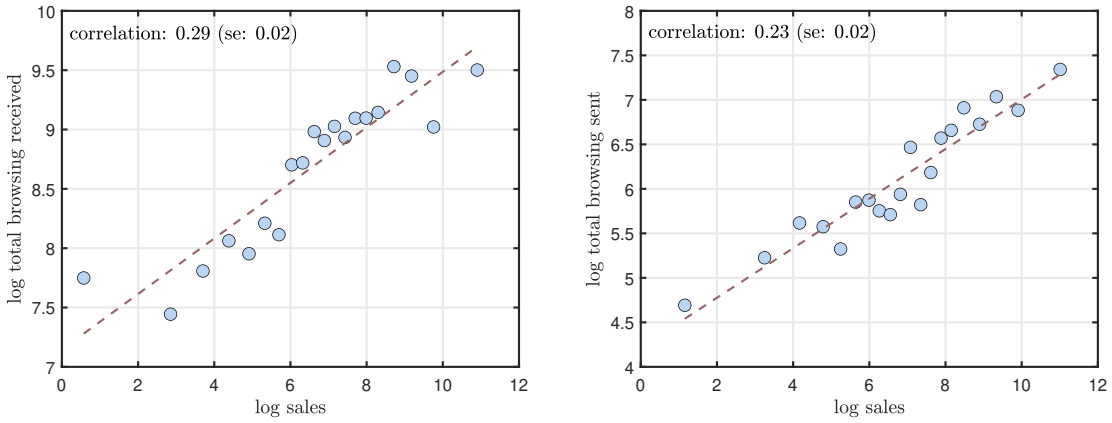
We next examine the relationship between firms' browsing activity and firm size, measured by sales. Using the 2016 sample, we find that both browsing received and sent are weakly positively correlated with firm size. The correlation between log browsing received and log sales is 0.29 (standard deviation:

[5] Using data in other years does not change our result.

[6] Gabaix and Ibragimov (2011) shows that the Rank-1/2 estimator reduces the bias of the OLS estimators.

0.02), while the correlation between log browsing sent and log sales is 0.22 (standard deviation: 0.02). Figure 3 presents the binscatter plots of these relationships, along with linear fits.

Figure 3: Browsing Received, Browsing Sent and Firm Size

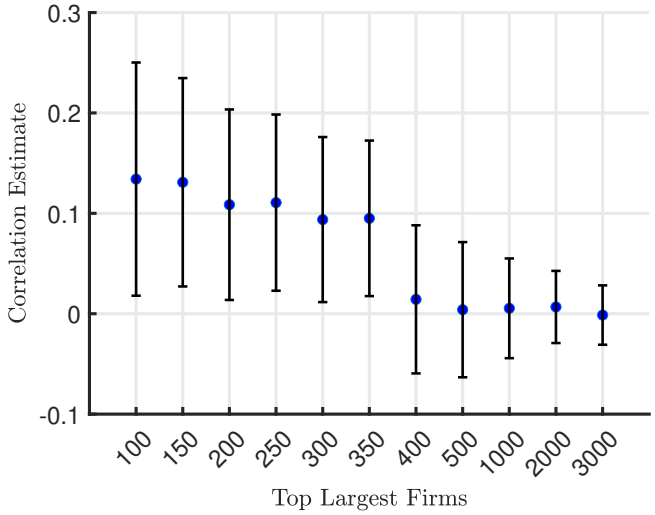


Fact 3. Browsing received and sent are positively correlated only for large firms

Finally, we investigate whether firms that browse others intensively also receive more attention. We find that this relationship holds only for large firms. Let R_i and S_i denote the total browsing received and sent by firm i, respectively. We compute the correlation $\text{corr}(R_i, S_i)$ for the top N largest firms, where N ranges from 100 to 3,000.

Figure 4 shows the estimated correlations and their 90% confidence intervals. A clear pattern emerges: the positive correlation between browsing received and sent is significant only for the top 350 largest firms, with a correlation of approximately 0.1. For smaller firms, there is no correlation.

Figure 4: Browsing Received-Sent Correlation and Firm Sizes



3. Framework

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

3.1 Environment

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

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

Household. The preference of the representative household is given by

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

where the parameter $\gamma > 0$ controls the income effects and $\nu \geq 0$ parameterizes the Frisch elasticity of labor supply. Here, ℓ_{it} is the labor supply of the workers who are sent to island i. The final goods consumption C_t is a composite of all the varieties, which admits a constant elasticity of substitution (CES) structure,

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

where ρ determines the elasticity of substitution across goods produced by different firms.[7] It follows that the aggregate price index and the corresponding demand schedule for a good i can be specified as

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

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

Finally, the budget constraint of the household is

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

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

Firms. The production technology of firm is given by

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

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

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

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

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

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

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

subject to

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

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

condition gives rise to the optimal labor demand condition

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

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

3.2 Best Response

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

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

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

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

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

where the coefficients are

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

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

To see exactly how an individual firm forms expectations about other firms' outputs, it is useful

to iteratate on condition (3.5) and reach the representation involves higher-order expectations

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

That is, firms need to form expectations on others' productivities, as well as expectations on other firms' expectations on productivities. and so on. Different from the envrionment with a continnum of agents Morris and Shin (2002) where agents only need to care about the economy average objects, firms in our model economy need to explicitly consider each firm's expectation due to the potential heterogeneity in information structure.

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

3.3 Information Network

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Bayesian updating formula, we obtain

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

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

4. Macro Elasticity to Micro Shocks

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

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

where g_i is the GDP elasticity with respect to firm i's idiosyncratic shock. These elasticities are results of firms' partial and general equilibrium considerations, as well as the expectations shaped by the information network. In this section, we provide the equilibrium characterization and shows that the magnitude of these elasticities hinge on an index of diluted attention received. We then provide the condition for granularity, where some firms' idiosyncratic shocks can impact the aggregate variables even when the number of firms goes to infinity.

4.1 GE and Higher-Order Expectations

Before providing the exact formula, we first illustrate at an intuitive level the three channels at work in determining the magnitude of g_i: (1) the efficiency channel: variations in z_{it} directly changes the production efficiency; (2) the information channel: z_{it} is informative about the aggregate component ξ_t and other firms' productivities; (3) the coordination channel: z_{it} of a salient firm i is more useful in forecasting other firms' decisions.

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

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

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

The second term in condition (4.1) involves the expectations about others' productivities. As

illustrated in subsection 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

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

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

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

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

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

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

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

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

where $\boldsymbol{m} = \begin{bmatrix} m_1 & \dots & m_N \end{bmatrix}$ *are the vector of in-degrees and* $\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 & \dots & \lambda_N \end{bmatrix}$ *are the vector of noise-to-signal ratios.*

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

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

The matrix $\mathbf{T}_1$ capture the effects directly observed productivities when forming expectations, and in-degrees $\mathbf{m}$ measure firms' salience. It corresponds to the first term of average first-order expectations

in . 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. It corresponds to the seocnd term of the average first-order expectations in ().

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

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

In this expansion, the first term $\frac{\delta}{N}\mathbf{I}$ correspond to the partial equilibrium effects that firms respond to their own productivity. The $k+1$-th term $(\mathbf{T}_1 + \mathbf{T}_2)^k$ then corresponds to the k-th order higher-order expectations in condition (4.1).

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

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

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

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

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

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

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

Though the magnitudes of g_i in these two cases are different, they share the property that when $N \to \infty$, $g_i \to 0$. That is, the idiosyncratic firm component has no aggregate impact when the number of firms goes to infinity. In these two special cases the information matrix $\mathbf{A}$ is symmetric, but can an asymmetric information matrix admits granular effect? In what follows, we answer this question next.

4.2 Bounds on the Elasticity g_i

As hinted in the previous discussion, in an information network, firms with a large in-degree tends to more important. Meanwhile, to which extent other firms are willing to react to a salient firm also depends on their own out-degrees, as observing many firms dilute their attention towards a particular salient firm. To gauge a firm's importance, we reach the following diluted in-dgreee index, which internalizes 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 understand this index, consider for example that all the firms who observe firm i do not observe any other firms. In this case, $d_j = 1$ for $j \in \mathcal{D}$,and the diluted in-degree index coincides with its original in-degree $\mu_i = m_i$. In contrast, as d_j becomes larger, the attention allocated to firm i becomes smaller, as measured by by $\frac{1}{d_j}$, which reminesent how it show the out-degree appears in the signal-to-noise ratio λ_i. It turns out that this diluted in-degree index is exactly what needed to pin-down the bounds on the elasticity g_i.

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

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

In principle, g_i relies on the entire information matrix, including information on both in-degrees and out-degrees. Proposition 4.2 shows that the diluted in-degree index μ_i succinctly summarizes these forces and is the key statistics to estimate an individual firm's quantitative importance.

To see how to obtain the bounds, one needs to leverage on the expansion (4.3). The lower bound include only the two terms—the shock itself and the first-order expectations—-and bounded it from below. In contrast, the upper bound include all the terms in the expansion and bound it from above.

A noteble feature of Proposition 4.2 is that it provides estimates for g_i without the need to solve the equilibrium. Computing these bounds only requires known structural parameters. Additionally, these bounds help understand the role of these structural parameters, particularly in quantitative analyses. For instance, reducing the relative volatility, as measured by $\tau_\xi/\tau_\varepsilon$, tends to increase the lower bound, as both the information and coordination channels are strengthened.

4.3 Granularity

In this section, we discuss the limiting case where the number of firms, N, approaches infinity. In standard macroeconomic models, the impact of an individual firm on aggregate GDP tends to diminish due to the law of large numbers. This effect can be seen in Proposition 4.2, as both the lower

and upper bounds are multiplied by $\frac{1}{N}$. This reasoning also applies to the two special cases involving perfect information or symmetric dispersed information. To allow for granular effects, it is necessary that, for some firms, the diluted in-degree index increases with N at a fast enough rate.

Proposition 4.3. *Then micro shocks can generate aggregate fluctuations if and only if* $\max_i \mu_i = \Theta(N)$.[8]

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

Corollary 4.1. 1. *The diluted in-degree index satisfies the following property*

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

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

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

Part 1 of Corollary 4.1 implies that mu_i is bounded from above by N, which only allows at most a small number (relative to the economy size N) of firms to have granular effects.

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

Example 1. In the first example, we assume that all firms to observe a fixed fraction ω of firms' productivities in the economy, in addition to their own. The left panel of Figure 5 displays the corresponding information matrix. Note that for the first ωN firms, they receive attention from the 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

$$d_i \approx \frac{1}{\omega N}.$$

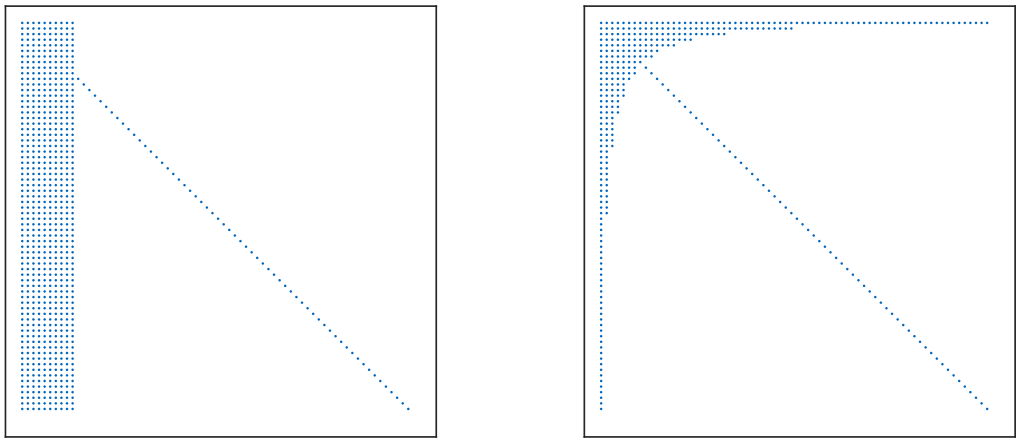
Can the first ωN firms exhibit granular effects? The answer is no. Notice that in this economy $\max_i \mu_i = \mu_1$, where

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

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

[8] 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



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

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

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

which is visualized by the right panel of Figure 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. Quantification

In the next section, we aim to quantify the extent to which micro shocks can induce aggregate fluctuations in an information network. We start with an extension that introduces additional heterogeneities which will interacts with the information heterogeneity. We then discuss the calibration procedure and the quantitative results.

5.1 Extension

Relative to the baseline specification, we allow heterogeneity in firm sizes and the volatility of the firm-specific compnenets. These changes lead to the following modifications. First, the best response function (3.5) and the GDP accounting equation become

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

where ω_i corresponding to firms' steady-state market share.

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

5.2 Calibration

Information Matrix. To connect the browsing activities from EDGAR and the information matrix in our framework, we adopt a random graph approach based on models of Chung et al. (2003b) and Chung et al. (2003a), which is a generalization of the classical Erdos-Renyi model of random graph. This random graph will assign probabilities for whether firm i can observe firm j' productivity.

Starting with the bilateral browsing matrix $\mathbf{B}$ constructed from EDGAR, we define $\widetilde{m}_i$ as the number of browsing request received by firm i and $\widetilde{d}_i$ as the number of browsing request sent by firm i in the sameple. In the model economy, A_{ij} can either be 0 or 1, and the probability that $A_{ij} = 1$ is given by

$$\mathbb{P}(A_{ij} = 1) = \min\left(c \left(\frac{\widetilde{m}_j}{\max_k \widetilde{m}_k} \right)^{\kappa} (1 + \varrho_j \widetilde{d}_i),\ \bar{P} \right). \tag{5.2}$$

In the formula (5.2), the term $\left(\frac{\widetilde{m}_j}{\max_k \widetilde{m}_k} \right)^{\kappa}$ captures the idea that if firm j receives more browsing requests in EDGAR, it is more likely that firm i observes j. The term $(1 + \varrho_j \widetilde{d}_i)$ captures the idea that if firm i sends more browsing requests in EDGAR, it is more likely that firm i will also observe firm j.

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

$j = 1, 2, ...N$,

$$\mathbb{E}\left[m_j\right] = \sum_{i=1}^{N} \mathbb{P}(A_{ij} = 1) \propto \left(\frac{\widetilde{m}_j}{\max_k \widetilde{m}_k}\right)^{\kappa_1}. \tag{5.3}$$

The parameters $\{\varrho_j\}$ controls for the correlations between in-degrees and out-degrees of the random graph. We employ the following formula,

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

and we discipline the parameter ϱ by tageting the degree correlations of the top 400-sales firms in the data and set $\varrho = 0.003$. Setting $\varrho_j = 0$ would render a firm's in-degree and out-degree independent.

Table 2: Calibrated Parameters

Param.	Value	Data Source	Empirical Target
Exogenously determined parameters			
γ	0.2	—	income elasticity
ν	2	—	Frisch elasticity
ρ	0.4	—	elasticity of substitution
θ	0.66	—	labor share
$\{\omega_i\}_{i=1}^N$	—	Compustat	sales' share
$\bar{\sigma}_i$	0.176	Compustat	weighted average of idiosyncratic productivity shock volatility
$\{\sigma_i\}_{i=1}^N$	—	Compustat	idiosyncratic productivity shock volatility
σ_ξ	0.02	San Francisco Fed	aggregate TFP shock volatility
Endogenously determined parameters			
c	0.2631	EDGAR	fraction of browser-browsee linkages
κ_1	1.8	EDGAR	power law exponent in the degree sequences of browsing matrix
ϱ	0.003	EDGAR	correlation between in-degree and out-degree in the browsing matrix
$\bar{P}$	0.95	EDGAR	maximal number of visitors for one single firm

We set the constant c in our model, which targets the fractions of information linkages (1s) existing in the data, that is, how many browser-browsee pairs exist in the data. We set $\bar{P} = 0.95$ to discipline the maximal in-degrees of the random graph so that it matches the maximal number of visitors one single firm has in the data. Together, the random graph model is completely characterized by quadruples $(\varrho, c, \kappa, \bar{P})$.

Structural Parameters. The structural parameters are calibrated based on standard values in literature. Specifically, we set the income elasticity $\gamma = 0.2$, the Frisch elasticity of labor supply to $\nu = 1$, the elasticity of substitution among different goods to $\rho = 0.4$, and the labor share $\theta = 0.66$. These structural parameters implies that the degree of strategic complementarity in our model is set to

$\alpha = 0.50$, while the response coefficient of output to productivity shocks is set to $\delta = 0.67$.

We calibrate firm size parameters and shock volatilities using Compustat and aggregate statistics for TFP from the San Francisco Fed.[9] Firm sizes are calculated as the average annual firm sales. To obtain shock volatility at the firm-level, we follow the standard approach in the literature to first estimate the firm-level annual TFP. We then calculate the standard deviation of annual TFP growth rate for each firm, as a measure of firms' shock volatilities. Averaging these volatilities using sales weight yields the average firm-level shock volatility.

5.3 Results

Table 3 summarizes the comparison across various model economies. For the information network, we consider three versions: the heterogeneous-information model corresponds our baseline calibration that targeted to the EDGAR browsing pattern; the perfect-information economy corresponds to the one with $A_{ij} = 1$ for all i and j; the Lucas-island economy corresponds to the cannonical dispersed-information economy with $A_{ij} = 1$ only when $i = j$. For all these economies, we start with the case where we mute the heterogeneity in firm sizes and shock volatilities, and we then allow these heterogeneities to play a role.

Table 3: Aggregate and Idiosyncratic Volatility

	Size and Volatility		
	equal size + equal volatility	hetero. size + equal volatility	hetero. size + hetero. volatility
Heterogeneous Info.			
total volatility	1.71	2.03	2.27
idiosyncratic volatility	0.49	0.92	1.00
idiosyncratic/total	28.37%	45.02%	44.10%
top 100 firms	0.44	0.86	0.97
top-100-sales firms	–	0.79	0.85
top-100-indegree firms	0.44	0.62	0.62
Perfect Info.			
total volatility	2.75	3.06	3.19
idiosyncratic volatility	0.39	1.42	1.66
idiosyncratic/total	14.13%	46.20%	52.16%
top 100 firms	0.06	1.34	1.61
top-100-sales firms	–	1.33	1.50
top-100-indegree firms	0.06	0.54	0.61
Lucas Island			
total volatility	1.36	1.52	1.60
idiosyncratic volatility	0.19	0.70	0.82
idiosyncratic/total	14.13%	46.31%	51.65%
top 100 firms	0.03	0.66	0.80
top-100-sales firms	–	0.66	0.75
top-100-indegree firms	0.03	0.27	0.30

Note: "—" indicates missing or unavailable data.

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

In the economy without heterogeneity in firm sizes and shock volatilities, we can isolate the role of information network. Due to the finite sample size (there are about 4000 firms in the sample), the law of large number does not apply exactly. That said, the share of GDP variance accounted by idiosyncratic shocks in the heterogeneous-information economy is much larger than the other two economies. The standard deviation of "top 100 firms" that are most important in accounting for GDP volatility is 0.44, which is around 7 times larger than that in the perfect-information economy and 15 times larger than that in the Lucas-island economy. There is no surprise that these important firms in the heterogeneous information economy are those "top-100-indegree firms", which receive most attention from the rest of the economy.

When firm size heterogeneity is allowed, the second column of Table 3 depicts a very different picture. With perfect information, micro shocks can account for a significant amount of GDP volatility, which is the firm-size channel proposed by Gabaix (2011). Most of the GDP fluctuations due to micro shocks are accounted by the top 100 firms and in this case, the majority of them are the top-100-sales firms. In our baseline heterogeneous-information economy, top-100-sales firms and top-100-indegree firms now play similar roles in contributing to aggregate fluctuations. The information channel remains quantitatively important. Moreover, the volatility of top-100-sales firms is smaller than that in the perfect-information economy, suggesting that the information channel actually dampens the firm-size channel.

Network Effects. To further understand these results, it is useful to revisit the discussion on higher-order expectations in subsection 4.1. In the extended model, the GDP can be expressed as

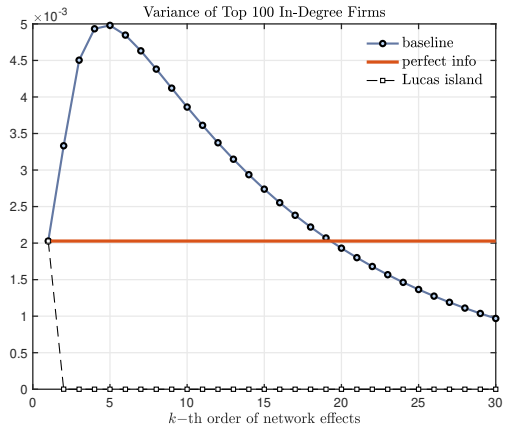
$$q_t = \underbrace{\delta \sum_i \omega_i z_{it}}_{\text{1st-order}} + \underbrace{\delta\alpha \sum_i \omega_i \mathbb{E}_{it}\left[\sum_j \omega_j z_{jt}\right]}_{\text{2nd-order}} + \underbrace{\delta\alpha^2 \sum_i \mathbb{E}_{it}\left[\sum_j \omega_j \mathbb{E}_{jt}\left[\sum_k \omega_k z_{kt}\right]\right]}_{\text{3rd-order}}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

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

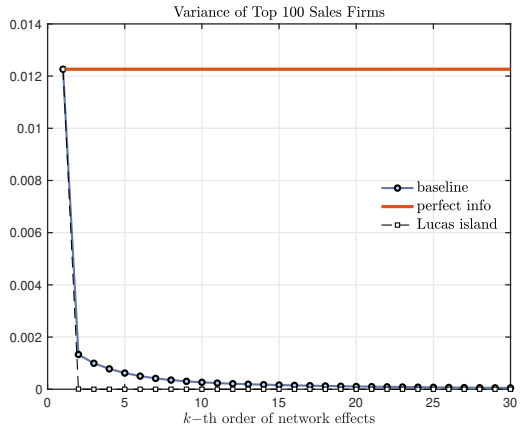
The left panel of Figure 6 displays the variance of k-th order network effects due to top-100 firms measured by their in-degrees. First notice that for the 1st-order term, it is common across all the economies as it is independent of the information network, which explains why the three economies start from the same point. In contrast, once the order increases, the information heterogeneity starts to play a role. With perfect information, higher-order and first-order expectations coincide with each other, and equals to the average productivity itself, results in the red line to be constant across different orders. In the Lucas-island economy, higher-order expectations are much more dampened

Figure 6: Role of Top Firms in Higher-Order Network Effects

(a) top-100 in-degree firms



(b) top-100 sales firms



than first-order expectations in responding to fundamentals, a key insight from Morris and Shin (2002), which explains the sharp decline of the dashed line after $k = 1$. In our baseline economy, the top-100 in-degree firms receive a lot of attention from other firms, and their productivities are similar to semi-public signals. If their productivities are true public signals, then we would expect the blue circular line to keep increasing as k increases, as public signals are more important in forming higher-order expectations. As semi-public signals, the variance of the top firms increases with k at the beginning as public signals, but eventually declines as k further increases. The increasing part provides a powerful amplification channel, which explains larger volatility accounted by top-100 indegree firm in Table 3.

The right panel of Figure 6 displays the variance of k-th order network effects due to top-100 firms measured by their sales share. Mechnically, due that that these firms are larger than top-100 in-degree firm, the network effect at $k = 1$ is larger. Unlike the left panel, the blue circular line monotonically declines with k, albeit at a slower rate than the Lucas-island economy. The reason is that not all the top-100 in-degree firms are large firms. In fact, only about 20% of the top-100 in-degree firms also belong to the top-100 sales firms. The relatively low correlation explains why the information channel actaully dampens reinforces the size-channel. These two channels only reinforce each other when the correlation between in-degrees and firm sizes are sufficiently positively correlated.

5.4 Testable Implications

Our model offers two novel predictions on 1) the relationship between firms' attention received and the comovment of their sales with GDP, and 2) the relationship between firms' attention received and the average correlation between a firm's forecast with others'. In this section, we show that these predictions are supported by the data.

5.4.1 Comovement with GDP and Attention Received

We assess if the sales of firms that receive more attention comove more strongly with GDP. To this end, we estimate the following regression equation,

$$\rho_i = \alpha + \beta_1 \log B_i^R + \beta_2 \log S_i + \beta_3 \log B_i^R \times \log S_i + \varepsilon_i, \tag{5.4}$$

where $\rho_i = \text{corr}(\Delta \log S_{it}, \Delta \log GDP_t)$ represents the correlation between firm i's quarterly sales growth and quarterly GDP growth, B_i is firm i's annual average total browsing received, and S_i is firm i's average sales.

Table 4 presents the baseline results. Column (1) shows that a 50 percent increase in a firm's browsing received leads to a 0.012 increase in the correlation between its quarterly sales with the quarterly GDP. Column (2) displays a relatively smaller magnitude when we replace firms' browsing with their sales — a 50 percent increase in a firms' average sales is associated a 0.0095 increase in the correlation. Column (3) shows that, even conditional on firms having the same sales, more browsing received is related to greater correlation between its sales with GDP. All regressions include standard firm controls and the industry fixed effect. We are interested in the interaction effect β_3, which reflects the complementary effect of firms' browsing received and sales in affecting the correlation. Column (4) presents the estimate of the correlation effect, which indicates significant complementary effects between browsing received and sales. All regressions include standard firm controls and industry fixed effect.

Table 4: Comovement with GDP, Sales, and Browsing Received

	Correlation with GDP			
	(1)	(2)	(3)	(4)
Log Sales		0.017***	0.013***	-0.0012
		(0.0020)	(0.0023)	(0.0075)
Log Browsing Received	0.024***		0.011***	0.0016
	(0.0033)		(0.0034)	(0.0042)
Browsing Received × Sales				0.0020**
				(0.00081)
Firm Controls	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓
No. Observations	9,321	9,320	9,320	9,320

Note: The industry fixed effect is at the 3-digit NAICS level and the standard errors (in parentheses) are clustered at the 3-digit NAICS level. The firm-level controls include the profit margin, leverage, operating leverage and ROA (return on assets). *Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5.4.2 Forecast Correlation and Attention Received

Our model predicts that the forecast of a firm, on average, have greater correlation with other firms' forecast, when this firm receives more attention. To test this prediction, we use the I/B/E/S Managerial Guidance dataset, which provides public companies' quantitative (numerical) managerial forecasts extracted from corporate earnings call transcripts and press release. In particular, a firm's manager makes forecasts on 14 variables related to his/her firm's business activities and performance, such as sales, earnings per share and capital expenditure etc. In our analysis, we use two variables: the EPS (earning per share) and sales, as they are most relevant to our topic and are surveyed most frequently, thereby containing the largest number of observations.

We then estimate the following regression equation on average forecast correlation,

$$\rho_i = \alpha + \beta_1 \log B_i^R + \beta_2 \log S_i + \beta_3 \log B_i^R \times \log S_i + \varepsilon_i, \tag{5.5}$$

where corr($Forecast_{it}, Forecast_{jt}$) represents the correlation between firm i's and firm j's forecast on their own sales/EPS in quarter t, so that ρ_i denotes the average correlation between firm i's forecast with the forecast of other firms. Here, B_i is firm i's annual average total browsing received, and S_i is firm i's average sales.

Table 5 presents the estimation results. Column (1) and Column (2) show that a firm's browsing received and sales both predict greater correlation between its forecast on its own sales with other firms' forecast on their own sales. The two estimates display similar magnitude — a 50 percentage increase in the browsing received or sales leads to an 0.013 increase of the correlation. Column (3) shows that the browsing received still significantly affects this correlation, even conditional on firms' sales. Results in Column (4), (5) and (6) on the EPS forecast display similar patterns. All regressions include industry fixed effect and standard firm controls such as leverage, profit margin and ROA.

Table 5: Correlation of Forecast, Sales, and Browsing Received

	Average Pairwise Forecast Correlation					
	Sales Forecast			EPS Forecast		
	(1)	(2)	(3)	(4)	(5)	(6)
Log Browsing Received	0.026***		0.012*	0.015***		0.0073*
	(0.0056)		(0.0073)	(0.0038)		(0.0040)
Log Sales		0.027***	0.021**		0.013***	0.0088**
		(0.0086)	(0.010)		(0.0036)	(0.0036)
Firm Controls	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓	✓	✓
R-squared	0.121	0.136	0.138	0.095	0.100	0.101
No. Observations	1268	1142	1142	1437	1365	1365

Note: The industry fixed effect is at the 3-digit NAICS level and the standard errors (in parentheses) are clustered at the 3-digit NAICS level. The firm-level controls include the profit margin, leverage, operating leverage and ROA (return on assets). *Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

6. Conclusion

In this paper, we document that firms' attention displays great heterogeneity. The attention firms receive follows a fat-tailed distribution based on their browsing activities related to each other's electronic filings. We propose a framework that combine the noisy business cycle framework with an information network. We provide the condition on the in-degree and out-degree of this information network that allows for granular effects. Quantitatively, the information channel is strong enough to generate sizable aggregate fluctuations driven by micro shocks. However, the information channel dampens the traditional firm-size channel as the correlation between firms size and received attention is not sufficiently high.

There are two promising directions for future work: first, in our framework, the information matrix is exogenously given. This is a useful benchmark for positive analysis, but less so for policy analysis. Extending the machinery of rational inattention to connect with our documented empirical regularities will be useful. Second, we omit the input-output linkages at the firm level. It will be interesting to see the interaction between the information network and the production network in regions where the firm-level transcation data are more avaiable.

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

A. Proofs of Results

A.1 Proof of Lemma 3.1

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

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

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

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

subject to

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

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

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

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

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

Note that in log-linear form,

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

Replacing ℓ_{it} by q_{it}, we have

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

where the two composite parameters are defined as

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

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

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

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

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

and the proof is now complete. □

A.2 Proof of Proposition 4.1

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

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

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

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

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

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

As such, the equilibrium conjecture follows

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

To derive the solution, it is necessary to use general representation (A.2). Using the in-out degree construction in

the main text, we obtain a compact representation for the expected aggregate output,

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

Next, we apply the inference formula (3.10),

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

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

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

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

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

Note here we apply the in-degree property,

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

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

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

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

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

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

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

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

We apply algebra tricks to simplify the last term,

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

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

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

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

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

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

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

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

A.3 Proof of Proposition 4.2

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

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

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

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

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

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

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

We also need the following definition of M-matrix.

Definition 2 (Plemmons (1977)). *An* $n \times n$ *matrix* $\mathbf{A}$ *is called an M-matrix if it can be written as* $A = k\mathbf{I}_n - \mathbf{B}$ *where* $\mathbf{I}_n$ *is the identity matrix, B is a non-negative matrix (every element of B is non-negative) whose spectral radius (the largest absolute value or complex modulus of its eigenvalues) is strictly less than k.*

The following are some properties of M-matrix.

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

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

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

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

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

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

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

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

Lemma A.2. *Let*

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

Then we have:

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

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

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

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

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

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

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

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

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

and lower bounded by

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

Since g_i *is equal to* $\frac{\delta}{N}$ *times the i-th row sum of* $(\mathbf{I} - \mathbf{V})^{-1} = \sum_{j=0}^{+\infty} \mathbf{V}^j$*, thus* g_i *has the following lower and upper bounds for* $1 \leq i \leq N$*, i.e.,*

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

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

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

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

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

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

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

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

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

Then by Theorem 1, we know that the spectral radius of $\mathbf{V}$ is at most α, thus by Definition 2 we obtain that $\mathbf{I} - \mathbf{V}$ is an M-matrix.

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

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

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

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

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

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

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

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

Then

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

Therefore,

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

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

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

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

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

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

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

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

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

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

However, the sum $v_1^{(j)}$ of elements in the first row of $\mathbf{V}^j$ is at most $j\alpha^j\cdot\mu_1$ by induction hypothesis. Also, notice that $\mathbf{v}_{11}\le\alpha$ and $\mathbf{v}_{1i}\le\frac{\alpha}{N}\cdot\mu_1$ for $i\ne1$, and by (4) we know the sum of all elements in $\mathbf{V}^j$ is at most $\alpha^j\cdot N$, thus the sum of elements in $\sum_{i=1}^N\mathbf{v}_{1i}\cdot\mathbf{V}_{i,j}^*$ (i.e., the sum $v_1^{(j+1)}$ of elements in the first row $\mathbf{V}^{j+1}$) satisfies

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

For other rows, the proof is the same.

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

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

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

Now we are ready to prove Proposition 4.2.

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

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

□

B. Additional Tables and Figures

Table 6: Filed Documents Being Viewed Most

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

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