Inattentive Network: Evidence and Theory*

Yue Fang
Zhejiang University

Zhen Huo
Yale University

Minghao Li
Peking University

Jieran Wu
Zhejiang University

August 20, 2024

Abstract

In this paper, we provide novel evidence on firms' attention allocation in a production network. Based on internet traffic log files collected by EDGAR, we construct a firm-level panel that measures firms' browsing intensity on other firms' electronically filed reports. We find that: (1) firms pay more attention to other firms that are closer to themselves in terms of network distance; (2) firms pay more attention to other firms that are more volatile; (3) a firm's absolute forecast error decreases in its total browsing activities. We then build a framework where firms rationally acquire information to set prices in a production network. The model's predictions on firms' attention allocation and price rigidities are consistent with the empirical patterns. In this framework, we show that the optimal monetary policy design significantly differs from that in a model where informational frictions are exogenously given, and the optimal policy endogenously induces additional dispersion in price rigidities.

Keywords: Rational inattention, production network, monetary policy

*We are grateful for comments from Costas Arkolakis, Olivier Coibion, John Leahy, Andrei Levchenko, Francesco Lippi, Chang Liu, Jianjun Miao, Daniela Vidart, Pengfei Wang, Michael Weber, and Yi Wen. The paper benefits from the discussion with seminar participants at CCERSI 2024, CICM 2024, DUFE, EFG conference on "Expectation, Information Frictions, and Economic Policy", Shanghai Jiaotong University, University of Connecticut, University of Michigan, TAMU, Peking University. We also thank Jinhui Xu for his excellent research assistance.

1. Introduction

In different sectors of an economy, firms display distinct pricing inertia while forecasting economic outcomes with various precisions. What economic forces drive these cross-sector heterogeneity in prices and information? How do firms allocate their limited attention in complex supply chains? How should a central bank implement the optimal monetary policy when sectoral nominal rigidities are interdependent and endogenous to their policy? These questions have become increasingly salient, given the emerging empirical evidence at both firm and industry levels.[1] Despite rapid development in recent literature that studies the macroeconomic impacts of information frictions, these questions remain largely unexplored.

This paper makes three contributions. Empirically, we provide a novel measure of firms' attention allocation based on their browsing history on others' electronic filings on EDGAR. We document that the intensity of the bilateral browsing decreases in their network distance but increases in the volatility of economic activities. Theoretically, we build a framework with rationally inattentive firms in a production network. We provide an analytical characterization of the equilibrium allocation as well as the optimal monetary policy rule. In particular, the optimal price stabilization index and the endogenous price rigidity interact with each other and are jointly determined. Quantitatively, the model yields properties of both the attention allocation and sectoral price rigidities that are consistent with data but absent with exogenous information. Due to the information acquisition channel, the optimal policy endogenously induces additional dispersion in price rigidities.

Attention measurement. Firms' attention allocations are crucial for their decision-making and for testing different theories on information acquisition, but they are not directly observable. To overcome this difficulty, we utilize the EDGAR "Logfile" which contains detailed information on the online viewing records for every SEC filing published on the EDGAR platform, including a partially masked IP address and the time of the access requested, among other information. We then uncover the browsers' complete IP addresses using the method developed in Chen et al. (2020), and map each viewer's uncovered IP address to their true identity with the service provided by ip-info.io. This procedure allows us to construct a panel of firms' browsing volume for other firms' files. The browsing intensities towards different firms, therefore, provide an account of attention allocation or, put it differently, an *attention network*.

We connect these browsing activities with firms' other economic activities and document the following facts. First, a firm pays more attention to others that are closer to itself in terms of network hierarchy. The amount of browsing declines significantly if a firm is not the browser's immediate supplier or customer based on the relationship information in the Factset data. In addition, the same logic extends to the dynamic setting: the browsing activity decreases (increases) after a relationship breaks (forms). Second, a firm pays more attention to others that are more payoff-relevant in terms of

[1] See Carvalho et al. (2021), Candia et al. (2023), and Pastén et al. (2024).

sales share and are more volatile. This observation also holds at the sectoral level, where the payoff relevance can be measured by the standard input-output linkages. Third, a firm's absolute forecast error decreases in its total browsing activities, consistent with the interpretation that more effort in information acquisition helps reduce informational frictions and improve forecast accuracy.

Theory. To rationalize the observed empirical patterns, we build a framework where firms in different sectors participate in a production network and have to post their prices without perfect information about sectoral productivity shocks. To overcome the informational frictions, firms can acquire any signals about various sectoral variables subject to a constant marginal cost in reducing uncertainty as in Sims (2003). When firms acquire a lot of information, their prices will be more responsive to underlying shocks, and the prices will be more flexible. In our environment, such endogenous information acquisition problem and the associated price rigidities are endogenous to both the network effects in general equilibrium and to the imposed monetary policy. As firms are connected via the production network, their information acquisition decisions strategically interact with each other. Meanwhile, the monetary policy directly affects how the nominal wage responds to sectoral shocks, which in turn will affect firms' payoff and be internalized in the information acquisition decision.

Despite these complications, we provide a compact characterization of the equilibrium price rigidity. Fixing a monetary policy, firms will only acquire a single noisy signal about their nominal marginal cost. The precision of this signal increases with the variance of the marginal cost process. The aforementioned network effects and policy effects are nested in the process of the marginal cost, and the equilibrium allocation boils down to a fixed-point problem between sectoral price flexibilities and the variance of marginal cost processes. We then provide a uniqueness and existence condition for this fixed-point problem. In terms of attention allocation, we measure firms' attention towards different sectors by the reduction in uncertainty of sectoral fundamentals. We show that firms in sector i allocate more attention to sector j if sector j is endowed with more volatile productivity shocks or if sector i's sectoral price is more exposed to sector j's shock in the influence matrix. This result is consistent with our empirical account of attention allocation and highlights the underlying (additional) general equilibrium forces.

Optimal monetary policy. Different from models in which rigidities are due to Calvo-type frictions or subject to exogenous informational frictions, the rigidities in our model are endogenous. This poses new challenges for the design of optimal monetary policy. The monetary policy now faces an additional expectation management role, which can influence firms' information choices. Furthermore, such management needs to take into account the fact that firms' information choices are also interconnected via the production network. We provide a closed-form optimal policy rule that takes the form of a price stabilization index. This price stabilization index contains two parts: the first exogenous part is holding firms' price rigidities unchanged when optimizing the monetary policy rule, which is identical to that in La'O and Tahbaz-Salehi (2022); the second endogenous part allows firms' price rigidities to vary with the policy, which is unique in our setting. The endogenous component succinctly summarizes the impact of information acquisition in shaping the optimal policy.

Notably, the optimal policy rule and the sectoral price flexibilities are jointly determined, which opens the door to policy-induced nominal rigidity. In a nutshell, the policymaker is incentivized to put more weight on more rigid sectors, which helps stabilize the prices in these sectors. However, the rigid sectors initially have less volatile marginal costs and are more reluctant to acquire information. This stabilizing effect from the monetary policy tends to reduce the volatility of marginal costs in these sectors, further dampening firms' incentive to acquire information and intensifying the price rigidity. This feedback channel between the optimal policy and endogenous information acquisition leads to additional dispersion of price rigidities relative to models where price rigidities are fixed ex-ante.

Quantification. To quantify these mechanisms, we calibrate the model to match the distribution of the absolute forecast error of earnings per share. The model also yields the distribution of price-change frequencies similar to the data. The calibrated model is able to produce the attention allocation pattern that is consistent with the sectoral evidence in the browsing activities. In addition, the model predicts a positive correlation between the sectoral shock volatility and the sectoral price-change frequency, which resembles the pattern in the data. Such correlation is natural in our model as firms tend to acquire more information when their payoff-relevant variables are more volatile. In contrast, it cannot be easily accounted for in models with ad hoc price rigidity.

We show that in the calibrated model, the optimal monetary policy rule significantly differs from that under exogenous information or with fixed information capacity. The difference in the policy rule can be almost entirely accounted for by the heterogeneity in the initial price rigidities. Due to the aforementioned policy-rigidity feedback channel, more rigid sectors will be assigned additional weight in the optimal policy with elastic information acquisition than that with exogenous information. At the same time, the sectoral price rigidities are much more dispersed once the optimal monetary policy is in place.

Related literature. Our paper complements an expanding literature that has already incorporated 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)). The subjects of these research span from aggregate fluctuations driven by sectoral public information, as in Nimark, Chahrour, and Pitschner 2019, to the interactions between higher-order beliefs and production networks, as in Bui et al. (2024). However, most of these papers focus on settings with exogenous information structures in which the degree of information (nominal) rigidities are taken as ad hoc model primitives. Instead, we introduce endogenous information acquisition of rationally inattentive firms, which aligns with the empirical evidence of firms' browsing activities and the attention network we establish.

We model firms' information acquisition as the optimal choice under the information-capacity constraint. This modeling choice places the paper in the vast literature of Rational Inattention (henceforth RI) initiated by the seminal contribution of Sims (2003, 2010).[2] Relative to this literature, our paper differs in two aspects. First, we extend the RI framework to a model that features both sectoral nominal rigidities and input-output linkages. Firms' optimal attention allocations are determined by their sectoral network positions. More importantly, attention choices among different sectors are interconnected via strategic complementarity. Afrouzi (2023) studies how strategic interactions (competition structure) within each sector affect attention choices and money non-neutrality. In contrast, we emphasize the role of underlying attention linkages across sectors in shaping the general equilibrium outcomes and the design of optimal policy. Second, we adopt the methodology of Miao et al. (2022) to solve for the unrestricted optimal information choices. The solution requires no prior assumption on the structure of signals, producing a compact representation of nominal rigidities as a function of firms' attentions. A contemporaneous paper by 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. With a pre-specified two-dimensional signal structure, the authors argue that attentions are centered on downstream firms due to their role as "information agglomerators".

In our model, the endogeneity of sectoral attention (nominal rigidity) allows for expectation management via monetary policy. In this regard, our paper is related to a small literature that discusses the design of optimal monetary policy with endogenous information frictions (Adam (2007); Paciello and Wiederholt (2013); Li and Wu (2016); Angeletos et al. (2020)). In previous papers, the optimal monetary policy is derived based on EITHER of the two notions of information endogeneity: (i) learning endogeneity in the sense that firms learn from prices (marginal costs), and (ii) attention endogeneity in the sense that firms' cognitive mistakes (noises) are determined by cost and benefit of information. We extend the optimal policy analysis to a network setting that captures both endogeneities, and we depart from previous literature by emphasizing the role of endogenous feedback between attention and policy in determining the optimal policy implementation.[3]

This paper is also related to the growing literature that examines the monetary transmission mechanism and designs optimal monetary policy in production networks.[4] The existing literature primarily focuses on models with time-dependent price settings such as Calvo frictions (e.g., Pasten et al. (2020), Ghassibe (2021), Rubbo (2023), and Pasten et al. (2024)). To the best of our knowledge, our paper is the first to study monetary transmission and optimal monetary policy in production

[2]Notable contributions include Maćkowiak and Wiederholt (2009, 2015), Caplin and Dean (2015), Matějka and McKay (2015), Luo et al. (2017), Caplin et al. (2018), Hébert and La'O (2023), Caplin et al. (2022), Angeletos and Sastry (2019), and Flynn and Sastry (2019).

[3]Ou et al. (2024) study the interactions between endogenous information acquisition and Calvo nominal rigidity in a two-sector model without input-output linkages. Abstracted from learning endogeneity in the information structure, their discussion of optimal policy is restricted to the class of price-stabilization indices. In contrast, our optimal policy analysis is unrestricted in the sense that the central bank can choose the money supply rule as an arbitrary function of productivity shocks.

[4]Previous literature focuses on optimal monetary policy in either horizontal or vertical economies, for example, Aoki (2001), Benigno (2004) and Huang and Liu (2005).

networks with state-dependent pricing. We demonstrate that monetary policy is able to shape the distribution of price stickiness in the production network. Enlightened by this feature, we propose a novel expectation management channel in the optimal monetary policy design.

2. Data and Browsing Activities

Exploiting a unique and novel data set that covers bilateral web browsing between firms, we document three new facts that offer insights on firms' attention allocation along the production networks:

Fact 1: A firm pays more attention to firms that are closer to itself in terms of network distance.

Fact 2: A firm pays more attention to firms that are more volatile and payoff-relevant.

Fact 3: A firm's absolute forecast error decreases in its total browsing activities.

2.1 Data and Stylized Facts

Our empirical analysis leverages three main sources of data: detailed firm-to-firm browsing data on documents filed at the EDGAR, survey data on firms' forecasts, and data on input-output linkages at both firm and industry levels. We begin by describing how we construct the browsing data, while offering stylized facts on firms' browsing activities. For more details, please refer to Online Appendix A.

2.1.1 The EDGAR Browsing Data

Our firm-to-firm browsing data is compiled from EDGAR, which stands for the Electronic Data Gathering, Analysis, and Retrival system. This system, maintained by the U.S. Secruties and Exchange Commision (SEC), operates as an electronic disclosure platform used by companies and other entities.[5] In particular, all U.S. public domestic companies are mandated to submit filings electronically on EDGAR since 1996. These filings, freely accessed through the internet, serve as a primary source to obtain firms' disclosed documents by, for example, firm managers, investors, and financial analysts.

The SEC's Division of Economic and Risk Analysis (DERA) started publishing data on internet traffic for all SEC filings in 2003. This dataset, often referred to as the "EDGAR Logfile," contains detailed information on the online viewing records for every SEC filing published on the EDGAR platform, including a partially masked IP address, the date and time of the access requested, and a unique identifier (the SEC assigns each document an "Accession number") for the disclosed document.

We use the term *browser* to refer to the viewer of a disclosed document, and *browsee* to refer to the owner of a disclosed document. To uncover the browsers' full IP address, we utilize the method

[5] The disclosures fall under the Securities Act of 1933, the Securities Exchange Act of 1934, the Trust Indenture Act of 1939, and the Investment Company Act of 1940. The SEC website provides more information on the EDGAR system: https://www.sec.gov/edgar/about.

outlined in Chen et al. (2020). We then use services provided by ip-info.io, a leading IP information provider, to map each viewer's uncovered IP address to their true identity. These steps result in a pairwise browers-browsee database at daily frequency covering the period from 2009 to 2016.[6]

We measure the *browsing intensity* of a browser i on a browsee j, denoted as b_{ijt}, by aggregating the number of online accesses from the browser to all filings owned by the browsee during a certain time interval t,

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

where a_{ijs} is equal to 1 if browser i requests an online access to browsee j's filings at time s, where time s is within the time interval t, and t can be a day, a month, a quarter or a year.[7] In the subsequent empirical analysis, we define the log browsing volume as $y_{ijt} = \log(b_{ijt})$.

Since attention is an internal cognitive process that can hardly be directly observed, measuring attention accurately can be a challenging task. We argue that the browsing intensity constructed above provides a reasonable measure of firms' attention allocation: we will observe more browsing records from firm i to firm j when firm i pays more attention to firm j.

The validity of using browsing intensity to quantify attention arises from two main aspects. First, firms' filings published on EDGAR contain the most timely and comprehensive information on their activities. For example, the SEC mandates firms to disclose unscheduled events through Form 8-K, ensuring prompt dissemination of crucial information. Additionally, Form 10-K, disclosed to the public only through EDGAR, offers annual updates in advance of the published annual reports, allowing critical business-related information to be released at an earlier stage.[8] For instance, General Electric's 2003 10-K was recorded 800 downloads from the company website, while downloads from the EDGAR amounted to 21,987 (4,325) during the year (two months) following its filing.[9] Second, although there exist other sources to access firms' filings such as Bloomberg or Yahoo Finance, EDGAR has become the primary source for obtaining firms' filings due to the following advantages: 1) all filings are freely available on EDGAR, while some are not on other platforms, and 2) critical information such as the income statement or balance sheets is often reported in pre-specified bins from other sources, leading to the loss of pivotal firm information compared to the original filings on EDGAR. As a result, EDGAR is likely to be viewed as the most important source of information when a firm needs to acquire information on other firms. Table B.2 in the Online Appendix lists the 10 types of filed documents with the most browsings.

---

[6] We select 2009 as our starting year because prior to 2009, the number of browsings are orders of magnitude less, due to significant changes in the disclosure requirement by the SEC and EDGAR. The SEC stops releasing SEC internet traffic data after 2017. More details are provided in the Online Appendix A.1.2.

[7] Time s can be measured as accurately as to a second. If t is set to, for example, 2015 January 2, b_{ijt} is equal to 3 when a browser requests online accesses to a browsee's document 3 times on that day.

[8] Note that 10-Ks are sometimes colloquially referred to as "annual reports". Here, "annual reports" refer to the document released online or in print shortly before the annual general meeting. The data show that annual reports lag 10-Ks for an average of 61 days.

[9] The article can be accessed on the Wall Street Journal website at https://www.wsj.com/articles/the-109-894-word-annual-report-1433203762.

The full sample consists of companies, institutions, and individuals. Our baseline sample is restricted to public firms for three main reasons. First, the browsees are primarily public companies, as only public companies are mandated by the SEC to file disclosures on EDGAR. Second, the successfully matched IP addresses are mostly owned by public companies. Finally, browsings of public companies can be matched with other firm-level data sources such as the Compustat, while information on private companies is typically unavailable. Our baseline sample contains a total of 713,157,510 unique IP addresses from 7,622 public companies, representing 52.4% of all browsings of disclosed files recorded in EDGAR during the sample period.[10] Online Appendix A.1.1 provides more details on how we construct our baseline sample.

Table 1 provides the summary statistics of our final baseline sample. 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.

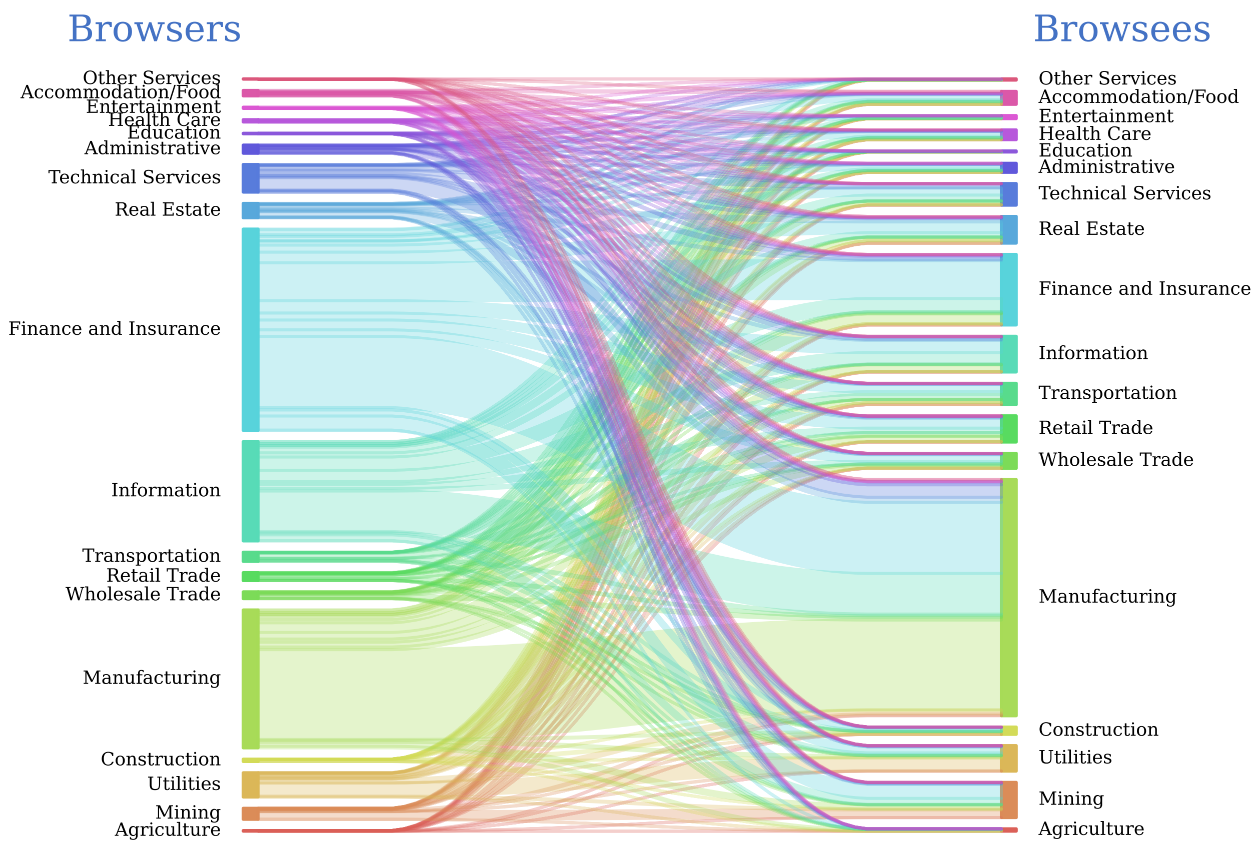
Table 1: EDGAR Browsing Data: Summary Statistics

	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

Finally, Figure 1 illustrates the bilateral browsing activities at the NAICS 2-digit industry level. Firms in the finance and insurance industry have the largest browsing volume on other firms, especially those in the manufacturing industry. Firms in the manufacturing industry also receive the most attention from the browsers. Online Appendix C provides additional facts on firms' browsing activities. In particular, Figure C.1 plots the distribution of total browsing volumes by firms and industries. Furthermore, we show that a firm browses other firms more intensely and also receives more browsings if it on average has larger sales and employees, or it is located in a less concentrated industry.

[10] It is noteworthy that automated crawling programs are becoming increasingly prevalent in recent years. To accurately measure firms' attention, we follow the method in Cao et al. (2023) to exclude IP addresses that browsed disclosures of over 50 companies in a single day or IP addresses that self-identified as bots in their user-agent headers. Our results are also robust to keeping those IP addresses in the sample.

Figure 1: Bilateral Browsing at the Industry Level



Note: This figure plots the bilateral browsing volumes between industries. The width of the bar is proportional to the browsing volumes. The media companies and financial investment firms are excluded.

The IBES Managerial Forecast Data. The I/B/E/S Managerial Guidance dataset provides public companies' quantitative (numerical) managerial forecasts extracted from corporate earnings call transcripts and press releases. Specifically, a firm's manager makes forecasts on the future realization of 14 variables related to his/her firm's business activities, such as capital expenditure, dividend per share, gross margin, etc. In our subsequent analysis, we select and 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 merge companies' "forecasts" of these two variables with the IBES Actuals, a dataset that documents the released realization of the corresponding variables, to calculate the forecast error used in our empirical analysis.[11]

The Factset Data. Our firm-level supply chain data is compiled using the Factset Revere-Supply Chain data from WRDS, a comprehensive database that provides detailed information on supply chain relationships disclosed by public firms, including the information of the supplying firm ("supplier")

[11] Companies sometimes only provide a forecast range with an upper and lower bound instead of an exact forecast value. In this case, we use the average of the upper and lower bound as the company's forecast.

and the procurer ("customer").[12]

We extract supply chain relationships in the dataset from 2009 to 2016, with the same time coverage as the browsing data. The original dataset records firm identifiers of the suppliers/customers, as well as contract start/end dates. In addition, approximately 11% of customer-supplier relationships report annual sales to the customer as a share of the supplier's total revenue that year. The processed dataset contains 4,260 unique suppliers and 4,518 customers. On average, each supplier reports 7.6 supply chain contract relationships, and each customer reports 7.55 such relationships in a year. Figure B.5 presents the number of firms and the forged supply chain relationships over time.

The BEA input-Ouput Database and Other Data Sources. The input-output linkages at the industry level are constructed using the input-output accounts published by the US Bureau of Economic Analysis (BEA). The BEA provides industry-level supply and use tables for 71 industries annually and 405 industries every five years. We transform the supply and use table to obtain the input-output (IO) table, an industry-by-industry matrix whose (i, j)th element represents industry i's input share from industry j. We obtain the annual IO tables for 71 industries from 1997 to 2021 and IO tables for 405 industries in 2007 and 2012. To link industries in the IO table with firms in other data sets, we employ the concordance between the BEA industry classification and the NAICS industry codes and map individual firms to BEA industries.

Our analyses make use of several other data sources. We use the Compustat dataset to obtain standard information on public firms and the firm-level TFP. We utilize the producer price index (PPI) data published by the Bureau of Labor Statistics (BLS) to calculate the industry-level inflation rates. We obtain the industry-level TFP from the BEA. The summary statistics for the IO table, industry-level inflation volatility, sales growth volatility, and TFP volatility are presented in Table B.1 in the Online Appendix.

2.2 Firm Level Evidence

This section provides empirical evidence at the firm level and offers two novel findings. First, a firm pays increasingly more attention to those firms that are closer to it —in the supply chain network sense. Second, a firm raises its attention to another firm when forming a new supplier-customer relationship. Conversely, a firm reduces its attention to another firm when they break an existing relationship.

2.2.1 Firms' Attention Hierarchy

To begin with, we convert the Factset sample into a year-supplier-customer format. Specifically, a supply chain linkage is considered to exist between firm i and firm j in year t if the date range of the

[12]The supply chain details in Factset Revere is obtained from public filings and annual/quarterly reports, transcripts of conference calls with investors and analysts, capital markets presentations, sell-side conferences, and firm press releases and websites, with public filings being the primary source (Culot et al. (2023)).

supply-chain relationship record overlaps with year t.

Next, we construct a network distance measure similar to the one in Carvalho et al. (2021). For each firm i in year t, all other firms are split into different groups. The "downstream distance 1" and "upstream distance 1" firms are the direct customers and suppliers of firm i, respectively. The "downstream distance 2" firms are customers of firm i's downstream distance 1 firms, but they are not distance 1 firms themselves. Following this recursive procedure, all firms are grouped based on their distance to firm i up to a distance of 4. Firms whose distance to firm i exceeds 4 are grouped together and serve as the control group.

To explore firms' attention along their supply chain networks, we estimate the following regression equation:

$$y_{ijt} = \alpha_{it} + \gamma_{jt} + \sum_{k=1}^{4} \beta_k^{\text{down}} \times \text{Downstream}_{ijt}^{(k)} + \sum_{k=1}^{4} \beta_k^{\text{up}} \times \text{Upstream}_{ijt}^{(k)} + \varepsilon_{ijt}, \tag{2.2}$$

where $\log(y_{ijt})$ is the log browsing volume of firm i on firm j in year t, while $\text{Downstream}_{ijt}^{(k)}$ and $\text{Upstream}_{ijt}^{(k)}$ are dummy variables indicating if firm j is firm i's downstream or upstream distance k firm, respectively, during year t. The terms α_{it} and γ_{jt} denote the firm i-year fixed effect and the firm j-year fixed effect. The coefficients of interest are β_k^{down} and β_k^{up}, which measure the differential browsing volume of firm i on firms with downstream and upstream network distance k relative to firms in the control group, respectively.

In Figure 2, the red dots and blue dots illustrate the regression coefficients β_k^{down} and β_k^{up}, measuring firms' attention hierarchy along the supply chain. Our results reveal two notable patterns. First, a firm tends to pay more attention to firms that are closer to it. On average, a firm pays about 60% more attention, as measured by EDGAR browsings, to distance 1 firms than to firms in the control group. Meanwhile, distance 2 firms receive roughly 10% more attention than those in the control group. The second pattern is that a firm allocates approximately the same amount of attention to its upstream suppliers and downstream customers within the same distance.

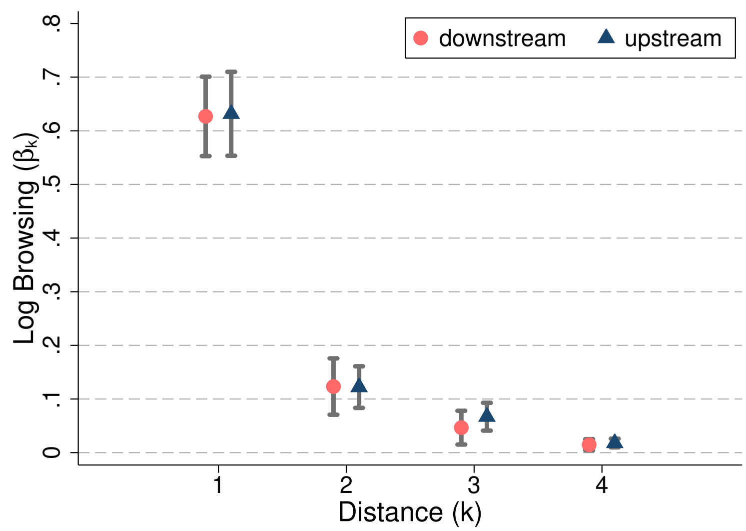
2.2.2 Sales Shares, Volatility, and Browsing Intensity

A subsample of firms in the Factset dataset provide information on their sales share for each customer. Leveraging this subsample, we further explore how a firm's attention paid to a customer varies with the firm's sales share to that customer and the customer's sales or TFP volatility. Specifically, we run the following regression equation:

$$y_{ijt} = \alpha_{it} + \beta_1 \text{sales\_share}_{ijt} + \beta_2 \sigma_{jt} + \text{controls}_{jt} + \varepsilon_{ijt}, \tag{2.3}$$

where $\log(y_{ijt})$ is the log browsing volume of firm i on firm j in year t, sales_share_{ijt} is defined as the share of firm i's sales to firm j as firm i's total revenue in year t, and σ_{jt} is the standard deviation of

Figure 2: Firms' Attention Hierarchy along the Supply Chain



firm j's sales growth/TFP over the past-5 years prior to year t. Here, α_{it} is the firm-year fixed effect, and ε_{ijt} is the error term. The controls consist of a broad array of time-varying covariates of firm j. The coefficients β_1 and β_2 measure how the browsing intensity of firm i depends on its sales share to firm j and the sales/TFP volatility of firm j, respectively.

Table 2 presents the estimates. Column (1) shows that a 10 percent increase in sales share to a customer leads to a 24.8 percent increase in the supplier's browsing volume to that customer, holding the customer's sales volatility constant. Conditional on the same sales share, when a customer's sales volatility increases by one standard deviation (equal to 0.16), the supplier's browsing volume increases by 7 percent. Column (4) displays a similar magnitude when we replace customers' sales volatility with the TFP volatility. When a customer's TFP volatility increases by one standard deviation (equal to 0.12), the supplier's browsing volume increases by 12.6 percent. Our results continue to hold when we add the time-varying controls and the firm-year fixed effect to the regression.

2.2.3 Firms' Attention when Forming/Breaking New/Existing Relationship: Event Studies

How does a firm allocate its attention when it enters a new trading relationship or ends an existing trading relationship with another firm? We exploit the panel feature of our data and employ event study methodologies to address this question. Specifically, we estimate the following regression equation:

$$y_{ijt} = \sum_{k=-m}^{n} \beta_k D_{ij(t-k)} + \alpha_i + \alpha_j + \gamma_t + \varepsilon_{ijt}, \tag{2.4}$$

where y_{ijt} is the log browsing volume of firm i on firm j in year t, and $D_{ij(t-k)}$ is a dummy variable, which is equal to one when a supplier and a customer established their trading relationship for the first time or terminated their relationship for the last time in the Factset dataset. At the same time,

Table 2: Browsing Intensity, Sales Share and Volatility (Firm Level)

	Browsing Intensity					
	(1)	(2)	(3)	(4)	(5)	(6)
Sales Share	2.48***	2.49***	1.72**	3.69***	3.79***	3.42
	(0.21)	(0.22)	(0.77)	(0.41)	(0.41)	(2.14)
Sales Volatility	0.44***	0.78***	0.83*			
	(0.16)	(0.19)	(0.48)			
TFP Volatility				1.05*	1.79***	1.33
				(0.54)	(0.62)	(1.23)
Controls		✓	✓		✓	✓
Firm-Year FE			✓			✓
Adjusted R^2	0.04	0.06	0.56	0.08	0.10	0.67
No. Observations	5154	4916	4916	1131	1106	1106

Note: This table shows how the browsing intensity of a firm depends on its sales share from the downstream customers, how volatile the customers' sales are, and how volatile the customers' TFP is. The control variables include firm size, age, leverage, fixed costs, stock returns, price-to-cost margins, and cash holdings. Robust standard errors are in parentheses and clustered by browsing firms. *Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

they did not have any trading relationships at least four years prior to the establishment or after the termination of the contract.

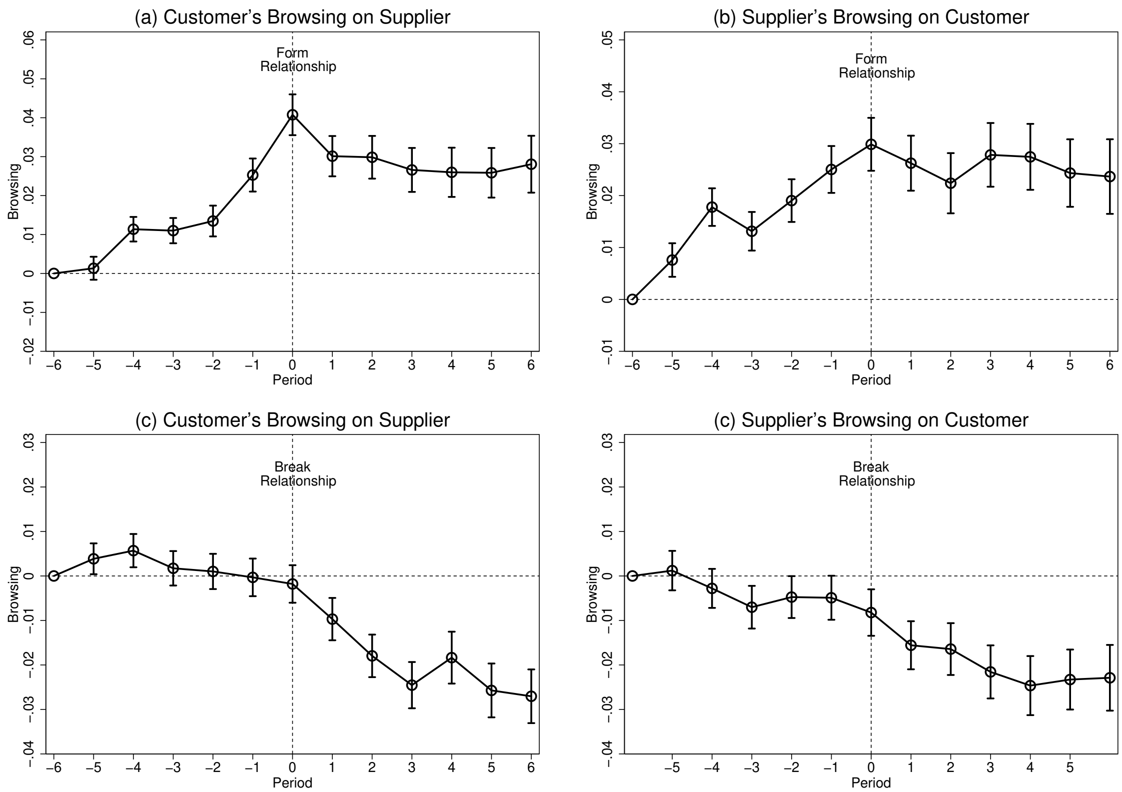
Figure 3 illustrates the estimated coefficients $\{\beta_k\}$.[13] Two patterns are notable. First, a firm gradually increases its attention toward its potential trading partner, reaching a peak in the quarter when the trading contract is signed. Afterward, its attention level remains steady. In contrast, a firm gradually reduces the attention it pays to the trading partner after the trading relationship ends. Second, a firm distributes roughly 3% more attention to its supplier or customer when their trading relationship is active.[14]

These results, together with our previous findings in Section 2.2.1, provide new insights into how firms allocate their attention. The primary factor determining firms' attention allocation is the distance between firms along the supply chain network. A firm prioritizes its attention towards potential (direct) suppliers or customers, even if it is currently not actively trading with them, while paying little attention to firms further down the supply chain. Despite the end of a contract, firms still closely monitor the activities of their potential suppliers and customers, maintaining almost the same level of attention as before.

[13]In the Online Appendix, Figure B.7 shows the estimates of the event when a supplier and a customer form their relationship for the first time or break it for the last time in the Factset dataset. Figure B.8 shows the estimates with 95 confidence intervals.

[14]We refer two firms to have an active trading relationship in quarter t when t falls within the date range of at least a trading contract recorded in the Factset data.

Figure 3: Browsing and Trading Relationship: Event Studies



Note: The figure depicts the log browsings of customers on suppliers (left panel) and suppliers on customers (right panel) when new trading relationships are formed at period 0. The frequency is quarterly and 90% confidence intervals are based on clustering at the browser level.

2.2.4 Browsing Intensity and Forecast Accuracy

Everything else equal, do firms forecast more accurately when their total browsings on other firms are greater? To assess this relationship, we adopt the following baseline empirical specification:

$$y_{it} = \alpha_i + \beta \sum_{k=t-n}^{t} browsing_{ik} + controls_{it} + \eta_t + \varepsilon_{it}.$$

Denote the variable being forecast as x_{iT}, whose actual value will be revealed at day T, and the forecast made by firm i at day t as $\widehat{x}_{it}$, where $t < T$. The forecast error is defined as $y_{it} = (\widehat{x}_{it} - x_{iT})/x_{iT}$, and the forecast horizon is defined as $T - t$. We focus on two scenarios where x_{iT} is either firms' earnings per share (EPS) or sales, as discussed in Section 2.1. The variable of interest is denoted as $\sum_{k=t-n}^{t} browsing_{ik}$, the log total browsing n days prior to day t when firm i makes the forecast. In our baseline specification, we set $n = 90$, approximately three months before firms making their forecasts. The firm fixed effect and time fixed effect are denoted as α_i and η_t, respectively.

Table 3 presents our baseline results. The dependent variable for the first three columns is firms' forecast error on their earning per share (EPS). Column (1) shows that when firms' browsing intensity increases by 50%, their forecast error reduces by 2.75 percentage points. This effect is substantial given that the median forecast error is 7.8%. Column (2) shows that this result continues to hold after controlling for a wide arrays of firm-level time-varying covariates. To address the concern that aggregate shocks could lead to simultaneous movements in firms' forecast error and their browsing intensity, column (3) adds time fixed effect to our regression. We see little change in the estimates. Finally, we include firm fixed effect to additionally control for unobserved time-invariant firm heterogeneity. The point estimate continues to remain economically and statistically significant. We may cautiously interpret this estimate as implying a causal relationship, as by adding the time and firm fixed effects, we are essentially demonstrating that a firm forecasts more accurately when it browses other firms more intensively 90 days prior to the forecast. In the Online Appendix Table B.4, we demonstrate that our results continue to hold when days prior to the forecast are set to 30 or 180. Estimates of forecast errors on sales display similar results, as shown in column (5)-(8), but with smaller magnitude.

Table 3: Browsing Intensity and Forecast Accuracy

	Forecast Error					
	Earnings Per Share			Sales		
	(1)	(2)	(3)	(4)	(5)	(6)
Browsing Intensity	-0.055***	-0.052***	-0.032**	-0.0020***	-0.0013*	-0.0017***
	(0.0082)	(0.0067)	(0.0097)	(0.00055)	(0.00057)	(0.00036)
Controls		✓	✓		✓	✓
Time FE		✓	✓		✓	✓
Firm FE			✓			✓
Adjusted R^2	0.01	0.10	0.25	0.04	0.15	0.42
No. Observations	7359	7076	6722	8329	7994	7607

Note: Standard errors are in parentheses and are clustered at industry level. This table shows how firms' forecast accuracy is associated with firms' browsing intensity in the past 90 days prior to making the forecast. The variables for firm controls include firm size, age, leverage, cash holdings, return on assets (ROA), fixed costs, stock returns, price cost margin and R&D intensity. *Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

2.3 Sectoral Level Evidence

We begin by examining the relationship between firms' browsing intensity and their input-output linkages. We show that a firm browses an industry more intensively when they allocate more expenditures to or gather more revenues from that industry. Firms also pay more attention to the industry with greater volatility of inflation, conditional on firms' expenditure share from and sales share to those industries.

We construct the annual industry-to-industry bilateral browsings by aggregating the daily firm-to-firm bilateral browsings. We then estimate the following regression equation:

$$y_{ijt} = \alpha_{it} + \beta_1 \text{input\_share}_{ijt} + \beta_2 \text{sales\_share}_{ijt} + \beta_3 \sigma_{jt} + \text{controls}_{jt} + \varepsilon_{ijt}, \tag{2.5}$$

where y_{ijt} is the log browsing volume of sector i on sector j at year t, and input_share$_{ijt}$ is defined as industry i's expenditure on goods produced by sector j as a fraction of its total expenditures on intermediate inputs in year t:

$$\text{input\_share}_{ijt} = \text{expenditure}_{ijt}/\text{total\_expenditure}_{it}.$$

Similarly, sales_share$_{ijt}$ is sector i's sales to sector j as a share of its total revenue, σ_{jt} is the standard deviation of sector j's annualized monthly inflation rate at year t, and ε_{ijt} is the error term. All regressions include the industry-year fixed effect α_{it}, which absorbs any industry-year-specific factors that affect industry i's browsing activity. Additionally, this fixed effect helps to address the possibility of spurious correlation due to industry-specific trends or common shocks to industries' browsings and the variables of interest. We further control for a list of potential industry j-year t confounding factors by adding time-varying covariates such as total revenues, total value added, and the share of intermediate inputs, etc. The coefficient β_1, β_2 and β_3 measure how the browsing intensity of firms in industry i depend on their input share from industry j, sales share to industry j and inflation volatility of industry j, respectively. These coefficients are likely to reflect conditional correlations rather than causal relationships, as firms' browsing volumes and the variables of interest are possibly endogenous, affected by other unobserved factors that we fail to control for.

Table 4 presents the main results. Column (1) to column (4) present results based on the annual 71-sector IO tables. Column (1) shows that a one standard deviation increase in the input share input_share$_{ijt}$, which is equal to 4 percentage points, is associated with 28 percent increase of industry i's browsing on industry j, when sector j is an upstream sector of sector i. This implies that, for example, firms browse 56% more intensively on industries at the 95th percentile of their input share than those industries at the 5th percentile. This number is 62.2% for the sales share, a similar magnitude as the input share.[15]. Column (5) to column (6) show coefficients with comparable (while a bit greater) magnitude for results based on the 405-sector IO tables.

We next shift our focus to estimating the relationship between a firm's browsing and the sales volatility of firms in its upstream and downstream sectors. Results in column (3) (column (8)) suggest that a one standard deviation increase in sector j's inflation volatility, corresponding to 9% (13%), leads to 25.6% (5.3%) increase in industry i's browsing volume on industry j for the 71-sector (405-sector) case. These results are little changed when industry i's input share and sales share on industry j are controlled for, as shown in column (4) and column (8). For robustness, in column (5) and column (10), we redefine the dependent variable as browser i's browsing volume on browsee j as a fraction of

[15]These positive relationships are illustrated in Figure B.1 in the Online Appendix

browser i's total browsing volume, $y_{ijt} = b_{ijt}/\sum_j b_{ijt}$, where b_{ijt} is given by equation (2.1). Our results are qualitatively unaffected.

In addition, Table B.3 in the Online Appendix shows that there exists positive relationship between firms' browsing volume and sectoral inflation volatility or sectoral TFP volatility of their upstream and downstream sectors. We estimate the same equation as in (2.5) except that σ_{jt} denotes inflation or TFP volatility.

Table 4: Browsing Intensity and Input-Output Linkages

	Browsing Intensity									
	71 Sectors					405 Sectors				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Input Share	7.00***			5.32***	0.34***	11.9***			8.98***	0.23***
	(0.97)			(0.96)	(0.75)	(0.84)			(0.79)	(0.02)
Sales Share		6.77***		4.99***	0.33***		11.5***		9.72***	0.34***
		(0.85)		(0.80)	(0.06)		(0.88)		(0.80)	(0.02)
Sales Volatility			2.84***	2.84***	0.027**			0.40***	0.41***	0.002***
			(0.26)	(0.23)	(0.013)			(0.034)	(0.032)	(0.0009)
Industry Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.79	0.79	0.79	0.81	0.18	0.67	0.67	0.67	0.68	0.23
No. Observations	26216	26216	26216	26216	25280	344298	344298	342296	309238	342176

Note: This table shows how the browsing intensity of an industry depends on its input share from the upstream sectors, the sales share from the downstream sectors, and how volatile the other sectors' inflation rates are. The control variables include average firm size, age, leverage, cash holdings, return on assets(ROA), fixed costs, stock returns, price cost margin, R&D intensity, Domar weight, total value added, total inputs and expenditure share of intermediate inputs. Robust standard errors are in paraentheses and clustered by industry. *Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Additional Evidence using Text-based Attention Measure. We supplement our empirical results in this section with the attention measure constructed using the Natural Language Processing (NLP) technique. Specifically, we compile a dictionary containing unique words that could best describe each industry at the 3-digit NAICS level. For each firm k in industry i in year t, we count how many times its annual 10-K document mentions an industry j, denoted as $a_{ijt}(k)$, based on the dictionary we have built. Aggregating the number of mentions within industry i, we obtain a measure of industry i's attention towards industry j, expressed as $a_{ijt} = \sum_{k\in i} a_{ijt}(k)$, and therefore construct an annual sequence of attention-allocation matrices across industries. Online Appendix A.2 provides more details on how we construct this measure.

We revisit our empirical results using this text-based attention measure and demonstrate that our findings continue to hold. In fact, the correlation between the text-based and the browsing-based attention measure is 0.32.[16] Table 5 shows that the results are consistent with our previous findings using browsing-based attention measure, both qualitatively and quantitatively.

Table 5: Text-based Attention and Input-Output Linkages

	Text-based Attention							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Input Share	3.07***			2.09***		2.15***		2.16***
	(0.42)			(0.42)		(0.41)		(0.40)
Sales Share		3.23***		2.51***		2.51***		2.49***
		(0.31)		(0.33)		(0.30)		(0.30)
Inflation Volatility			1.02***	0.98***				
			(0.049)	(0.047)				
Sales Volatility					2.12***	2.10***		
					(0.073)	(0.069)		
TFP Volatility							4.32***	4.26***
							(0.20)	(0.17)
Industry Controls	✓	✓	✓	✓	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.77	0.77	0.77	0.78	0.77	0.78	0.77	0.78
No. Observations	36684	36684	32347	32347	36684	36684	36684	36684

Note: This table shows how an industry's text-based attention depends on its input share from the upstream sectors, the sales share from the downstream sectors, and how volatile the other sectors' inflation rates, sales and TFP are. Robust standard errors are in paraentheses and clustered by industry. *Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3. Theory

In this section, we present the baseline model that connects with the empirical findings. Our framework builds on La'O and Tahbaz-Salehi (2022), and there is an important difference: firms actively choose their signals and the information structure is endogenously determined. We will characterize how firms' attention allocation and aggregate outcomes are jointly determined in equilibrium and how they are shaped by different policy rules.

Throughout, we use small letter x_t to denote a variable X_t's log-deviation from its steady state.

[16] To calculate this correlation, at the industry level within each browser-year cell, we first calculate the correlation between the text-based and browsing-based attention measure. We then calculate the average over browser-year cells.

3.1 Setup

Firms. The economy is comprised of N sectors indexed by i and j, and each sector consists of a continuum of firms indexed by ι. Firms in the same sector operate the same constant returns to scale production technology

$$Y_{i,\iota,t} = Z_{i,t} L_{i,\iota,t}^{\alpha_i} \prod_{j=1}^{N} X_{i,\iota,j,t}^{a_{ij}},$$

where $Z_{i,t}$ is the technology shock in sector i

$$z_{i,t} \equiv \log Z_{i,t} \sim \mathcal{N}(0, \sigma_i^2),$$

$L_{i,\iota,t}$ is the labor input, and $X_{i,\iota,j,t}$ is the intermediate input from sector j. The parameter α_i measures the relative importance of labor, and a_{ij} measures the relative importance of intermediate inputs. The production network is therefore captured by the input-output matrix with elements defined by a_{ij}. The covariance matrix of sectoral shocks is diagonal: $\mathbf{\Sigma}_z = \text{diag}\left(\{\sigma_i^2\}_{i=1}^N\right) \succ 0$.

Within a sector, each firm ι produces a different variety. The aggregate goods in sector i is given by the standard Dixit-Stiglitz aggregator with a constant elasticity of substitution θ_i, which yields a downward-sloping demand schedule

$$Y_{i,t} = \left[\int Y_{i,\iota,t}^{\frac{\theta_i - 1}{\theta_i}} d\iota\right]^{\frac{\theta_i}{\theta_i - 1}}, \qquad Y_{i,\iota,t} = \left(\frac{P_{i,\iota,t}}{P_{i,t}}\right)^{-\theta_i} Y_{i,t}.$$

In each period, there are two stages. In the first stage, firms post their prices potentially under incomplete information about the realization of sectoral technology shocks. They can acquire additional information beyond their prior, but this is a costly process. In the second stage, all the shocks realize and information becomes public. Firms then make their hiring and intermediate inputs decision to cater for the demand.

In the second stage, given the wage rate W_t and the vector of sectoral goods price $\{P_{jt}\}$, the nominal marginal cost of a firm in sector i is

$$\text{MC}_{it} = \frac{1}{Z_{it}} W_t^{\alpha_i} \prod_{j=1}^{N} P_{jt}^{a_{ij}}.$$

In the first stage, anticipating the marginal cost process and the demand schedule, firms' pricing and information acquisition problem is

$$\max_{P_{i,\iota,t}, \boldsymbol{x}_{i,\iota,t}} \quad \mathbb{E}\left[\left(\frac{P_{i,\iota,t}}{P_{i,t}}\right)^{-\theta_i} Y_{i,t}\Big((1+\tau_i)P_{i,\iota,t} - \text{MC}_{i,t}\Big) \,\Bigg|\, \boldsymbol{x}_{i,\iota,t}\right] - \mathcal{F}_i(\boldsymbol{x}_{i,\iota,t}). \tag{3.1}$$

where $\tau_i = \frac{1}{\theta_i - 1}$ is a standard subsidy rate that eliminates the price markup at the steady state.

Crucially, when setting the price $P_{i,\iota,t}$, firms do not observe the sectoral productivity shocks yet and therefore they have to form expectations about the marginal cost. Meanwhile, firms are rationally inattentive. They can actively acquire additional signals $\boldsymbol{x}_{i,\iota,t}$ subject to some information acquisition costs denoted by $\mathcal{F}_i(\boldsymbol{x}_{i,\iota,t})$, in the spirit of Sims (2003). In Subsection 3.2, we will specify the details of the cost function.

Crucially, as the information acquisition decision is determined in equilibrium, the implied price rigidity is endogenous to the general equilibrium effects and to the government policy. With incomplete information, firms' pricing decision will differ from that under perfect information. The allocation in the second stage may not be efficient due to that the average price level may underreact or overreact to shocks and that the cross-sectional relative price movements may be distorted. This leaves room for policy intervention via the aforementioned endogenous information acquisition channel.

Households. There is a representative consumer who maximizes her utility under perfect information

$$\max_{C_t, L_t} \frac{C_t^{1-\gamma}}{1-\gamma} - \frac{L_t^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}},$$

subject to

$$P_t C_t \leq W_t L_t + \Pi_t + T_t, \quad \text{and} \quad P_t C_t \leq M_t.$$

In the consumers' problem, γ determines the income effects and η corresponds to the Frisch elasticity. Consumers total expenditure has to be financed by the sum of labor income, firms' total profit (Π_t), and government transfer (T_t). In addition, consumers also need to respect the Cash-in-Advance (CIA) constraint: their total nominal expenditure cannot be larger than the exogenous money supply M_t. The final consumption goods C_t and the final goods price P_t are given by the following aggregators:

$$C_t = \prod_i C_{it}^{\beta_i}, \qquad P_t = \prod_{i=1}^{N} P_{it}^{\beta_i},$$

where β_i represents the expenditure share in goods from sector i.

Taking prices and wage rate as given, the optimal labor supply condition is given by

$$\frac{W_t}{P_t} C_t^{-\gamma} = L_t^{\frac{1}{\eta}}. \tag{3.2}$$

Monetary policy. We assume that the monetary authority is not subject to informtional frictions and can commit to the following monetary policy rule

$$m_t = \log M_t = \sum_{j=1}^{N} \psi_j z_{j,t}.$$

After prices have been set and all shocks have realized, the nominal wage will adjust one-to-one with the money supply m_t so that the CIA constraint is satisfied. Therefore, by varying the responsiveness to sectoral shocks, the policymaker can influence the processes of sectoral marginal costs, which in turn affects firms' endogenous price response in equilibrium.

3.2 Endogenous Information Acquisition

Frictionless benchmark. Consider momentarily the frictionless benchmark, in which case there is no information acquisition cost and firms can perfectly observe all the underlying shocks. The solution to firms' problem (3.1) then takes a particular simple form: the sectoral price is identical to the nominal marginal cost $P_{i,\iota,t} = P_{it} = \mathrm{MC}_{it}$. In terms of log-deviation from steady state, the marginal cost in sector i is given by

$$\mathrm{mc}_{it} = -z_{it} + \alpha_i w_t + \sum_j a_{ij} p_{jt}. \tag{3.3}$$

Solving for the vector of sectoral prices thus leads to the following characterization.

Proposition 3.1. *In absence of information acquisition cost,*

1. *The equilibrium price responses relies on the standard Leontief inverse matrix*

$$\boldsymbol{p}_t^* = \mathbf{L}(-\boldsymbol{z}_t + \boldsymbol{\alpha} w_t), \quad \textit{with} \quad \mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1}, \tag{3.4}$$

   where the (i,j)-th element of $\mathbf{A}$ is a_{ij} and $\boldsymbol{\alpha} \equiv \left[\alpha_1, \ldots, \alpha_N\right]'$.

2. *The equilibrium output change is independent of money supply*

$$c_t^* = \frac{1+\eta}{\gamma+\eta} \sum_i \lambda_i z_{i,t},$$

   where $\lambda_i = \sum_j L_{ij} \beta_j$ is the steady-state Domar weight of sector i.

This proposition provides the ideal price response function under perfect information. The familiar Leontief inverse summarizes the impact of production network on firms' price setting. As expected, when prices are flexible, the monetary policy is irrelevant and the standard Hulten's theorem applies in determining the real allocation.

Elastic attention. Now we specify the details of firms' information acquisition problem. Instead of observing the nominal marginal cost perfectly, firms need to form expectation about it. The nominal marginal costs depend on the vector of sectoral productivity shocks $\boldsymbol{z}_t$, and firms have to allocate limited capacity in learning about these underlying states. In our baseline analysis, we adopt the elastic attention approach (Maćkowiak et al., 2023), where the marginal cost in reducing uncertainty measured in mutual information is constant. The following lemma first transforms the original problem (3.1) into a tracking problem.

Lemma 3.1. *Under a quadratic approximation, firms' problem with elastic attention can be written as*

$$\max_{p_{i,\iota,t},\mathbf{H}_i,\mathbf{V}_i} \quad -\frac{1}{2}\lambda_i\theta_i\mathbb{E}\left[\left(p_{i,\iota,t}-mc_{it}\right)^2\right]-\chi_i\mathbb{I}\left((\boldsymbol{z}_t;\boldsymbol{x}_{i\iota t}|\Sigma_{\boldsymbol{z}}\right)$$

subject to

$$\begin{aligned} p_{i,\iota,t} &= \mathbb{E}[mc_{it}|\boldsymbol{x}_{i,\iota,t}], \\ \mathbb{I}\left(\boldsymbol{z}_t;\boldsymbol{x}_{i,\iota,t}\right) &= \frac{1}{2}\left(\log\det\boldsymbol{\Sigma}_z-\log\det\boldsymbol{\Sigma}_{z|x}\right), \\ \boldsymbol{x}_{i,\iota,t} &= \mathbf{H}_i\boldsymbol{z}_t+\boldsymbol{u}_{i,\iota,t}, \qquad \boldsymbol{u}_{i,\iota,t}\sim\mathbb{N}(\mathbf{0},\mathbf{V}_i). \end{aligned}$$

Lemma 3.1 illustrates the basic tradeoff faced by firms when choosing the information set: more accurate signals reduce the pricing error, but it incurs additional information acquisition cost. Here, λ_i is sector i's steady-state Domar weight which controls the benefit of additional information, and χ_i is the constant marginal cost when reducing mutual information between states and signals $\mathbb{I}\left((\boldsymbol{z}_t;\boldsymbol{x}_{i\iota t}|\Sigma_{\boldsymbol{z}}\right)$.

In principle, firms can choose multiple signals that contains arbitrary combinations of the underlying states and idiosyncratic noises (parameterized by $\mathbf{H}_i$ and $\mathbf{V}_i$). However, based on the technique developed in Miao et al. (2022), the optimal signal structure takes a rather simple form, as shown in the the following proposition.

Proposition 3.2. *For any linear laws of motion of the aggregate variables, the optimal signal structure satisfies*

1. *When $\chi_i<\lambda_i\theta_i\mathbb{V}(mc_{it})$, where $\mathbb{V}(mc_{it})$ denotes the volatility of sector i's marginal cost, firms acquire a single signal about the marginal cost*

$$x_{i,\iota,t}=mc_{it}+u_{i,\iota,t}, \qquad u_{i,\iota,t}\sim\mathbb{N}(0,\nu_i^2)$$

where $\mathtt{mc}_{it}$ satisfies condition (3.3) and the variance of the private noise ν_i^2 is given by

$$\nu_i^2=\frac{\chi_i\mathbb{V}(mc_{it})}{\lambda_i\theta_i\mathbb{V}(mc_{it})-\chi_i}.$$

2. *When $\chi_i\geq\lambda_i\theta_i\mathbb{V}(mc_{it})$, firms acquire no new information.*

The optimal signal structure is quite intuitive: firms choose to obtain a noisy version of their ideal price change and the noise level is increasing in the marginal cost χ_i. A couple of remarks are in order. First, the laws of motion of the nominal marginal costs as a function of the underlying states $\boldsymbol{z}_t$ are equilibrium objects, taken as given by individual firms with rational expectations. Consequently, the extent to which firms learn about different sector's productivity hinges on the processes of marginal costs. Second, the precision of the signal is increasing in the volatility of the marginal costs, and therefore, are also endogenous to the general equilibrium and monetary policy rule.

With the optimal signal structure derived in Proposition 3.2, firms set their prices according to the expected marginal costs $p_{i,\iota,t} = \mathbb{E}[\mathrm{mc}_{it}|x_{i,\iota,t}]$. Standard Bayesian inference directly implies the following pricing formula after averaging out idiosyncratic noises.

Corollary 3.1. *At the sectoral level, the optimal pricing strategy with elastic attention is given by*

$$p_{i,t} = \int p_{i,\iota,t} d\iota = \mu_i \, mc_{it}, \quad where \quad \mu_i \equiv 1 - \frac{\chi_i}{\theta_i \lambda_i \mathbb{V}(mc_{it})} \in [0,1]. \tag{3.5}$$

The variable μ_i shapes the price rigidity: a higher μ_i implies greater responsiveness to variations in nominal marginal costs and therefore smaller nominal rigidity. Note that μ_i inherits the property of the precision of optimal signals, and is increasing in the variance of the nominal marginal costs as well.[17] This property is in contrast with some commonly used alternative information structure which we discuss next.

Alternative information acquisition specifications. We consider two alternative specifications for information acquisition. The first one is the case firms face a fixed information capacity constraint when collecting information (Sims, 2003; Maćkowiak and Wiederholt, 2009). This corresponds to setting $\chi_i = 0$ in Lemma 3.1 while requiring

$$\mathbb{I}(\boldsymbol{z}_t; \boldsymbol{x}_{i,\iota,t}) \leq \delta_i,$$

where δ_i is some exogenous parameter that determines the maximum capacity a firm can choose to reduce faced uncertainty.

The second alternative is the exogenous information approach where the signal structure and its precision are exogenously determined. Particularly, we consider the specification as in La'O and Tahbaz-Salehi (2022), in which firms in sector i observe a vector of signals about $\boldsymbol{z}_t$

$$x_{i,\iota,j,t} = z_{j,t} + u_{i,\iota,j,t}, \qquad u_{i,\iota,t} \sim \mathbb{N}(0, \tau_i \sigma_j^2).$$

Notice that when inferring productivities in different sectors, firms face the same signal-to-noise ratio controlled by the exogenous parameter τ_i.

It turns out that both of these two alternatives yield observational equivalent pricing strategy as our baseline elastic attention case.

Proposition 3.3. *With alternative information structures, the optimal pricing strategy is*

$$p_{i,t} = \mu_i \, mc_{it},$$

[17] In Proposition F.1 of Appendix F, we formally establish the equivalence between firms' total attention and the nominal rigidity μ_i.

where $\mu_i = 1 - e^{-2\delta_i}$ with fixed information capacity and $\mu_i = \frac{1}{1+\tau_i}$ with exogenous information.[18]

It is worth noting that though the pricing formula in these alternatives resembles that from the elastic attention approach, the implied price rigidity μ_i in both cases are determined only by exogenously parameters and are independent of equilibrium outcomes or the monetary policy. In Section 5, we will provide further evidence that helps select the relevant approach in acquiring information.

3.3 Equilibrium

We start with the definition of equilibrium in the linearized economy.

Definition 1. *A competitive equilibrium consists a policy rule for the money supply m_t, the wage rate w_t, the sectoral price vector $\boldsymbol{p}_t = \left[p_{1,t}, \ldots, p_{N,t}\right]'$, and the consumption and labor allocation $\{c_t, \ell_t\}$, such that*

1. *Consumers' optimal labor supply condition (3.2) is satisfied.*
2. *CIA constraint is satisfied: $c_t + \sum_i \beta_i p_{i,t} = m_t$*
3. *Firms choose prices and information optimally so that condition (3.5) is satisfied.*

The key to understand the equilibrium outcomes is firms' pricing strategy. Once sectoral prices are determined, the total output c_t can be obtained via the CIA constraint. Thanks to Corollary 3.1, the optimal signal structure permits a compact characterization of the influence matrix.

Proposition 3.4. *Given a law of motion of the nominal wage $w_t = \boldsymbol{\kappa} \boldsymbol{z}_t = \sum_j \kappa_j z_{j,t}$, the price rigidities $\{\mu_i\}_{i=1}^N$ solve the following fixed-point problem*

$$\mu_i = \mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa}) = 1 - \frac{\chi_i}{\theta_i \lambda_i \mathbb{V}(mc_{it})}, \qquad \text{and} \qquad \mathbb{V}(mc_{i,t}) = \left\| \mathbf{e}_i (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} (-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa}) \boldsymbol{\Sigma}_z^{\frac{1}{2}} \right\|^2, \tag{3.6}$$

where $\boldsymbol{\mu} = \mathrm{diag}(\mu_1, \ldots, \mu_N)$ is a diagonal matrix of nominal rigidities and $\mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa})$ is attention best response function for firms in sector i. $\mathbf{e}_i$ denotes ith standard basis (row) vector in $\mathbb{R}^N$.

Proposition 3.4 highlights the endogeneity of price rigidities. Firms' choice of μ_i not only depends on the volatilities of underlying shocks ($\boldsymbol{\Sigma}_z$), but also depends on other firms' information acquisition decision and the wage function. Intuitively, when firms in sector j acquire more information, their prices become more responsive to underlying shocks, which in turn increases the volatility of the marginal costs via production network linkages. As a result, it provides incentive for firms in sector i to also acquire more information. In contrast, when firms face fixed information capacity constraint or when information is exogenous, μ_i becomes constant and the aforementioned feedback effect is

[18] In this paper we adopt natural logarithm log in computing the Shannon entropy, and the unit of information is called a "nat." The usual base for logarithm in the entropy formula is 2 , in which case the unit of information is a "bit." Adopting natural logs simplifies the algebra without changing any of our results.

muted. Meanwhile, the process of marginal cost also depends on how the nominal wage responds to different shocks, which translates into firms' information acquisition decision. These considerations are summarized in the best response function $\mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa})$, an object we will revisit when exploring the optimal policy design.[19]

Given the price rigidities $\boldsymbol{\mu}$ and the wage function $\boldsymbol{\kappa}$, the sectoral prices and the aggregate output gap can be derived accordingly.

Corollary 3.2. *1. The influence matrix of prices ϕ is*

$$\boldsymbol{p}_t = \boldsymbol{\phi} \boldsymbol{z}_t = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1} \boldsymbol{\mu}(-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa})\boldsymbol{z}_t. \tag{3.7}$$

2. The output gap is given by

$$c_t - c_t^* = -\frac{\eta}{1 + \gamma\eta} \sum_i \beta_i e_{i,t}, \tag{3.8}$$

where $e_{it} \equiv p_{i,t} - p_{i,t}^*$ is the average pricing error in sector i relative to the perfect-information benchmark

$$\boldsymbol{e}_t = \mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right) \boldsymbol{z}_t. \tag{3.9}$$

where $\mathbf{Q} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu})$, and $\mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1}$ is the Leontief inverse with perfect information

Relative to the frictionless benchmark (3.4), the influence matrix of prices (3.7) is akin to a modified Leontief inverse where the network matrix $\mathbf{A}$ is dampened by $\boldsymbol{\mu}$. This dampening matrix generates both sluggish response of prices and across-sectors distortion, which leaves room for policy intervention.

The highly-nonlinear equilibrium system characterized in Proposition 3.4 and 3.6 does not admits closed-form solution in general. Additionally, multiple equilibria is also a pervasive phenomenon in this types of models with endogenous information acquisition and strategic complementarity (Hellwig and Veldkamp, 2009). Equilibrium multiplicity invalidates the use of comparative statics methods to analyze the impact of parameter perturbations on equilibrium outcomes; it also complicates welfare and policy analysis. To address this problem, we present the following proposition that guarantees equilibrium uniqueness.

Proposition 3.5. *There exists a unique fixed point $\boldsymbol{\mu}$ that satisfies condition (3.6) if for each sector $i = 1, 2, ...N$,*

1. The monetary policy accommodate a wage rule that satisfies $\kappa_i < 1$;

[19] In Appendix F, we present a special example with closed-form characterizations, which further clarifies the dependence of sectoral nominal rigidities and endogenous feedbacks of attentions on model primitives $(\mathbf{A}, \boldsymbol{\Sigma}_{\mathbf{z}}, \boldsymbol{\kappa})$. In particular, we show that a sector's nominal rigidity (attention) is increasing in its shock volatility and its relative importance as a supplier —- two salient predictions that are captured in our quantitative analysis.

2. *The information cost* χ_i *satisfies* $0 < \chi_i < \theta_i \lambda_i \min\left\{\sigma_i^2 (1-\varkappa_i \kappa_i)^2,\ \ \sigma_i^4 (1-\varkappa_i\kappa_i)^4 (2\varsigma_i)^{-1}\right\}$. *Denote* $\boldsymbol{\Gamma} = -\mathbf{I} + \boldsymbol{\alpha\kappa}$, *and the auxiliary parameters* $\{\varkappa_i\}$ *and* $\{\varsigma_i\}$ *are defined as*

$$\varkappa_i = \begin{cases} 1, & 0 \le \kappa_i < 1 \\ \alpha_i, & k_i < 0 \end{cases}, \quad \textit{and} \quad \varsigma_i = \mathrm{tr}\left(\mathbf{L}\left|\boldsymbol{\Gamma}\boldsymbol{\Sigma}_z\boldsymbol{\Gamma}'\right|\mathbf{L}'\mathbf{e}_i'\mathbf{e}_i\mathbf{L}\mathbf{A}\right).$$

Proposition 3.5 provides a set of sufficient conditions that ensures equilibrium existence and uniqueness. The proof of this result utilizes the Kellogg's Fixed Point Theorem in Banach spaces (Kellogg, 1976). [20] We believe that technical strategy we adopt to derive these results has its independent value in methodology. In particular, it can be applied to other high-dimensional models of endogenous information frictions when establishing existence and uniqueness properties.

Finally, we provide the condition that relates the wage rate with the money supply.

Proposition 3.6. *In an equilibrium, the monetary supply and the wage rate jointly satisfy*

$$m_t = \frac{\eta}{1+\gamma\eta} w_t + \left(1 - \frac{\eta}{1+\gamma\eta}\right) \sum_i \beta_i p_{i,t} + \frac{1}{1+\gamma\eta} \sum_i \lambda_i z_{i,t}.$$

That is, in an equilibrium, the monetary authority can alternatively chooses the wage function $\boldsymbol{\kappa}$, obtain the price function via the mapping (3.6), and derive the corresponding monetary policy rule afterwards. This is the strategy we follow when solving for the optimal policy.

3.4 Attention Allocation

In this subsection, we discuss how firms in equilibrium allocate their limited attention to different sectors. In Section 2, we measure firms' attention on different sectors by their browsing activities. To connect with this moment in our theoretical framework, we first define the attention allocation to different sectors as a normalized reduction in uncertainty about sectoral fundamentals.

Definition 2. *Define the attention from firms in sector* i*'s on productivity in sector* j *as*

$$\omega_{ij} = \frac{\sigma_j^2 - \widehat{\sigma}_{j|i}^2}{\sigma_j^2}, \tag{3.10}$$

where $\widehat{\sigma}_{j|i}^2 = \mathbb{E}\left[\left(\mathbb{E}[z_{j,t}|x_{i,\iota,t}] - z_{j,t}\right)^2 |x_{i,\iota,t}\right]$ *is the posterior variance conditional on the chosen signals.*

An increase of ω_{ij} implies that firms in sector i is better informed about sector j's productivity. That is, more attention is allocated to learning about sector j's conditions. Recall that firms' optimally choose their signal structure to be the nominal marginal costs with noises. As a result, to which extent their signals reveal the fundamentals about the underlying states in different sectors depends on the

[20] In Appendix D, we also provide a more general and less restrictive conditions that can be used in our quantitative and policy analysis.

endogenous exposure of marginal costs to shocks in different sectors. The following proposition highlights how such attention allocation hinges on both the primitive shocks and equilibrium forces.

Proposition 3.7. *The attention allocation satisfies*

$$\omega_{ij} = \mu_i \frac{\sigma_j^2 \phi_{ij}^2}{\sum_k \sigma_k^2 \phi_{ik}^2} \qquad \textit{and} \qquad \sum_j \omega_{ij} = \mu_i,$$

where ϕ_{ij} is the element of the influence matrix $\boldsymbol{\phi}$.

To unpack this proposition, first note that the total attention in sector i equals to the price responsiveness μ_i. Second, to understand the allocation of attention across sectors, we leverage the equilibrium condition (3.5) that relates prices and marginal costs and the price influence matrix

$$\mathrm{mc}_{it} = \frac{1}{\mu_i} p_{it} = \frac{1}{\mu_i} \sum_j \phi_{ij} z_{jt}.$$

Recall that the marginal cost is driven by firms' own productivity shocks, the wage rate, and prices in different sectors, which are functions of the underlying states $\boldsymbol{z}_t$. In equilibrium, these forces are summarized by the influence matrix $\boldsymbol{\phi}$ and the informativeness of firms' signals about sector j's condition is determined by the volatility of $\phi_{ij} z_{jt}$. A higher volatility of z_{jt} therefore amplifies its relative importance in firms' signals, which shifts the attention to sector j. Meanwhile, ϕ_{ij} captures the general equilibrium exposure of $\mathtt{mc}_{it}$ to z_{jt}. When such exposure intensifies, the relative importance of sector j's component endogenously looms larger.

From the monetary authority's perspective, the design of monetary policy has an impact on both a firm's total attention acquired as well as the relative attention allocation across sectors. In the next section, we characterize the optimal monetary policy rule when internalizing its effects on firms' attention and implied price rigidity.

4. Optimal Monetary Policy with Endogenous Attention

In Section 3, we characterize how equilibrium sectoral attentions are determined by network input-output linkages and the volatilities of sectoral shocks. We also establish the attention linkages among sectors driven by strategic complementarity in information acquisition. In light of these results, in this section we study the optimal monetary policy in our model with endogenous attention and expectations.

4.1 Central Bank's Optimization Problem

The welfare loss in our model economy is caused by price errors due to incomplete information. Similar to La'O and Tahbaz-Salehi (2022), the following lemma decomposes the welfare loss into three components which are different functions of price errors.

Lemma 4.1. *The second-order approximation to the expected welfare loss is*

$$L = \frac{1}{2}\left[\left(\gamma + \frac{1}{\eta}\right)\mathbb{V}(c_t - c_t^*) + \sum_{i=0}^{N}\lambda_i \mathcal{C}_i + \sum_{i=1}^{N}\lambda_i\theta_i\mathcal{D}_i\right] \tag{4.1}$$

where the cross-sector price error dispersion is

$$\mathcal{C}_i = \mathbb{E}\left[\sum_{j=1} a_{ij}e_{jt}^2 - \left(\sum_{j=1}^{n} a_{ij}e_{jt}\right)^2\right] \qquad \mathcal{C}_0 = \mathbb{E}\left[\sum_{j=1}^{n}\beta_j e_{jt}^2 - \left(\sum_{j=1}^{n}\beta_j e_{jt}\right)^2\right],$$

and the within-sector price dispersion $\mathcal{D}_i$ is

$$\mathcal{D}_i = \mathbb{E}\left[\int_0^1 (p_{i,\iota,t} - p_{it})^2 d\iota\right] = \mu_i^2 \nu_i^2.$$

The first component of welfare loss corresponds to the variance of the output gap. This part is only related to a weighted average of sectoral pricing errors according to condition (3.8) Note that eliminating the output gap volatility only requires the average pricing error remains constant, $\sum_j \beta_j e_{jt} = 0$. This is the insight developed in Rubbo.

The second component corresponds to the cross-sector price error dispersion. Even when the aggregate output gap volatility is muted, the sector-specific pricing errors can still be present. These pricing errors will in turn lead to relative movements among different sectors that are inefficient from the planner's perspective.

The third component corresponds to the within-sector price dispersion. Even when the average prices at the sectoral level match those without informational frictions, idiosyncratic noises in firms' signals still generate inefficient cross-sectional dispersion. The magnitude of such dispersion depends on both the variance of idiosyncratic noises ν_i^2 and firms' responsiveness to signals μ_i. In our environment, these two objectives are jointly determined and ν_i^2 can be expressed as a function of μ_i^2 via condition (3.5).

According to condition (3.9), the sectoral pricing errors are functions of the price rigidity $\boldsymbol{\mu}$ and the wage function $\boldsymbol{\kappa}$, which is a property that also inherited by the output gap volatility and the cross-sector price error dispersion in the social welfare loss function. Also recall that for each wage function, there exists a monetary policy rule that supports it as an equilibrium. As a result, it is sufficient to consider the following policy design problem.

Lemma 4.2. *The optimal policy solves the following problem*

$$\min_{\boldsymbol{k}} L = \frac{1}{2}\left[\left(\gamma + \frac{1}{\eta}\right)\boldsymbol{\beta}'\boldsymbol{\Sigma}_e\boldsymbol{\beta} + \boldsymbol{\lambda}'\operatorname{diag}(\boldsymbol{\Sigma}_e) - \boldsymbol{\lambda}'\operatorname{diag}\left(\mathbf{A}\boldsymbol{\Sigma}_e\mathbf{A}'\right) - \boldsymbol{\beta}'\boldsymbol{\Sigma}_e\boldsymbol{\beta} + \boldsymbol{\chi}'\boldsymbol{\mu}^v\right],$$

where $\mathbf{\Sigma}_e = \mathbf{Q}(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})\mathbf{\Sigma}_z(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})'\mathbf{Q}' \geq 0$ *is the covariance matrix of the cross-sectional average of sectoral pricing errors, and* $\boldsymbol{\mu}$ *solves the fixed point problem defined in Proposition 3.4,* $\boldsymbol{\mu}^v = \mathrm{diag}(\boldsymbol{\mu})$ *is the vector of sectoral nominal rigidities.*

This optimal policy problem in our model incorporates two novel ingredients, relative to the existing literature. First, price flexibilities $\boldsymbol{\mu}$ are endogenous to the general equilibrium of the choice of the monetary policy, the policymaker possesses an extra policy dimension of influencing the equilibrium attention/price flexibilities, which creates policy room for managing expectations. This channel is absent in models with exogenous information frictions where price rigidities are exogenous and invariant to the policy choice.

Second, the expectation management interacts with input-output linkages. The attention choices in different sectors are interdependent through the structure of the marginal cost processes. The policymaker faces a non-trivial problem of designating price flexibilities across sectors, taking into account the relation between the sectoral distribution of attentions and the production network.

4.2 Optimal Policy with Expectation Management

Conceptually, with elastic attention, variations in the policy instrument $\boldsymbol{\kappa}$ influence the social welfare via two channels: the first is the exogenous channel when holding firms' attentions/rigidity $\boldsymbol{\mu}$ constant, and the second is the endogenous attention channel allowing responses of $\boldsymbol{\mu}$ to policy changes. The first-order condition with respect to $\boldsymbol{\kappa}$ naturally yields this decomposition

$$\frac{dL}{d\boldsymbol{\kappa}} = \underbrace{\frac{\partial L}{\partial \boldsymbol{\kappa}}}_{\text{exogenous chanel}} + \underbrace{\frac{\partial L}{\partial \boldsymbol{\mu}}\frac{d\boldsymbol{\mu}}{d\boldsymbol{\kappa}}}_{\text{attention channel}} = \mathbf{0}. \tag{4.2}$$

With exogenous information, fixed information capacity, or Calvo type pricing frictions, only the exogenous channel $\frac{\partial L}{\partial \boldsymbol{\kappa}}$ is at work as the implied nominal rigidity is invariant to policy changes or general equilibrium. What is unique in our setting is that the choice of μ_i in each sector responds to both the monetary policy $\boldsymbol{\kappa}$ and to other sectors' attention choices $\{\mu_j\}$ via general equilibrium forces. In what follows, we will spell out the details of the attention response $\frac{d\boldsymbol{\mu}}{d\boldsymbol{\kappa}}$, the welfare exposure $\frac{\partial L}{\partial \boldsymbol{\mu}}$, and finally the formula of the optional monetary policy.

Exogenous channel. As a benchmark, we first discuss the case where only the exogenous channel is at work. In fact, due to the equivalence result provided in Proposition 3.3, setting $\frac{\partial L}{\partial \boldsymbol{\kappa}} = \mathbf{0}$ leads to the same optimal monetary policy as that in an environment with exogenous information or fixed information capacity. The following proposition provides the characterization of such optimal policy.

Proposition 4.1. *With fixed information capacity or exogenous information, the optimal monetary policy is* *identical and can be implemented via a price stabilization policy*

$$\sum_{i=1} \varphi_i^x p_{it} = 0, \tag{4.3}$$

where the sectoral weights given by

$$\varphi_i^x = \Bigg[\underbrace{\frac{(1-\rho_0)}{(\gamma+1/\eta)}\lambda_i}_{\text{output gap}} \;+\; \underbrace{\sum_{j=1}^{N} (1-\mu_i)\,\lambda_j \rho_j l_{ji} + (\rho_0 - \rho_i)\,\lambda_i}_{\text{cross-sector dispersion}} \;+\; \underbrace{\mu_i \lambda_i \theta_i \rho_i}_{\text{within-sector dispersion}} \Bigg] \left(\frac{1}{\mu_i} - 1\right), \tag{4.4}$$

and ρ_i for $i = 1, \ldots, N$ and ρ_0 are given by

$$\rho_i = \mathbf{e}_i (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} \boldsymbol{\alpha}, \quad \rho_0 = \sum_i \beta_i \mu_i \rho_i.$$

This optimal policy rule is identical to that in La'O and Tahbaz-Salehi (2022) with exogenous signals, and it turns out that the same rule also applies to the economy with fixed information capacity. While the signal structures in these two economies are substantially different, they are nevertheless equivalent in the eyes of the policymaker. The three terms in (4.4) addresses the three different inefficiencies in the policymaker's objective function (4.1), respectively.

The weight on each sector not only depends on the Domar weight, but also depends on its interaction between sectoral rigidities captured by ρ_i. In La'O and Tahbaz-Salehi (2022), ρ_i is interpreted as the upstream price flexibility of sector i, and this measure will also show up in the endogenous component of the policy rule when allowing the attention channel.

Attention response to policy change. Importantly, when deriving the formula (4.4), sectoral attentions or nominal rigidities $\{\mu_i\}$ are taken as exogenously determined. Now we allow them to be chosen by firms. Recall from Proposition 3.4, firms attention choices are characterized by a fixed-point problem, and their best response $\mathcal{T}(\boldsymbol{\mu}, \boldsymbol{\kappa})$ hinges on all other firms' decisions and the wage function. A change in the wage function directly translates into changes in the marginal cost processes in different sectors. Due to the network linkages, firms' attention choices are also interconnected and have to be jointly determined. The following proposition formalizes these considerations.

Proposition 4.2. *Given the attention best response $\mathcal{T}(\boldsymbol{\mu}, \boldsymbol{\kappa})$, the impact of changes in monetary policy $\boldsymbol{\kappa}$ on equilibrium attention $\boldsymbol{\mu}$[21] are given by an $N \times N$ matrix*

$$\frac{d\boldsymbol{\mu}^v}{d\boldsymbol{\kappa}} = \left[\mathbf{I} - \mathcal{T}_{\boldsymbol{\mu}^v}\right]^{-1} \mathcal{T}_{\boldsymbol{\kappa}}, \tag{4.5}$$

[21]Since $\boldsymbol{\mu}$ is a diagonal matrix, we focus on its diagonal vector to ease the algebra of matrix calculus.

where the associated Jacobian matrices of derivatives admit the following representations

$$\mathcal{T}_{\boldsymbol{\kappa}} = 2\,\mathrm{diag}\left\{\frac{1-\mu_i}{\mathbb{V}(mc_{it})}\right\}\left[\mathbb{COV}(\boldsymbol{mc}_t, \boldsymbol{z}_t) \odot \left((\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1}\boldsymbol{\alpha}\mathbf{1}'\right)\right], \tag{4.6}$$

$$\mathcal{T}_{\boldsymbol{\mu}^v} = 2\,\mathrm{diag}\left\{\frac{1-\mu_i}{\mathbb{V}(mc_{it})}\right\}\left[\mathbb{COV}(\boldsymbol{mc}_t, \boldsymbol{mc}_t) \odot \left((\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1}\mathbf{A}\right)\right], \tag{4.7}$$

where $\odot$ denotes the Hadamard product and $\mathbb{COV}$ denotes the covariance operator.

In Proposition 4.2, the (i, j) elements of matrices $\mathcal{T}_k$ and $\mathcal{T}_{\boldsymbol{\mu}^v}$ measure the marginal effects of changing policy instrument κ_j and sector j's attention μ_j on an individual firm's attention in sector i, respectively. The matrix $\left[\mathbf{I} - \mathcal{T}_{\boldsymbol{\mu}^v}\right]^{-1}\mathcal{T}_{\boldsymbol{\kappa}}$ then encodes all the general equilibrium consideration when allowing firms to internalize that all other firms' attention choice will further adjust according to their attention best response functions.

To unpack the results, consider first the (i, j) element of $\mathcal{T}_k$, which is given by

$$\frac{\partial \mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa})}{\partial \kappa_j} = \frac{1-\mu_i}{\mathbb{V}(\mathrm{mc}_{it})}\mathrm{Cov}(\mathrm{mc}_{it}, z_{jt})\rho_i. \tag{4.8}$$

This formula summarizes how a different wage response to z_{jt} affects firms' attention choice μ_i via altering the variance of the marginal cost. The first term, $\frac{1-\mu_i}{\mathbb{V}(\mathrm{mc}_{it})}$, captures the variability of sector i's attention. As μ_i approaches 1, it becomes increasingly difficult to influence its attention choice. The second term, $\mathrm{Cov}(\mathrm{mc}_{it}, z_{jt})$, corresponds to the exposure of the marginal cost to the sectoral productivity shock. The last term, $\rho_i = \mathbf{e}_i(\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}\boldsymbol{\alpha}$, measures the importance of such wage change from a network perspective. Note that even holding other firms $\boldsymbol{\mu}$ unchanged, a different wage function still implies different price dynamics in different sectors, while ρ_i in turn summarizes the consideration of these changes for a firm in sector i.

In the same vein, the (i, j) element of $\mathcal{T}_{\boldsymbol{\mu}^v}$ can be written as

$$\frac{\partial \mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa})}{\partial \mu_j} = \frac{1-\mu_i}{\mathbb{V}(\mathrm{mc}_{it})}\mathrm{Cov}(\mathrm{mc}_{it}, \mathrm{mc}_{jt})h_{ij}. \tag{4.9}$$

Different from the case for κ_j, a marginal change in μ_j modifies the response of firms in sector j to all shocks. The exposure to such change for firms in sector i is effectively the covariance between their marginal costs, $\mathrm{Cov}(\mathrm{mc}_{it}, \mathrm{mc}_{jt})$. Since the price dynamics of all other sectors will also adjust given that p_{jt} is different, firms in sector i also need to adjust accordingly. Parallel to the last term in condition (4.8), the importance of such network effects is measured by h_{ij}, which is the (i, j) element of $(\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}\mathbf{A}$.

So far, the discussion of conditions (4.8) and (4.9) remains at the partial equilibrium level, as it holds other firms' attention choices unchanged. In general equilibrium, once sector i's attention choice changes, it will induce endogenous response of all other sectors' attention choices, which in turn will lead to another round of attention modification, and so on. This infinite order of general equilibrium reasoning can be summarized by the matrix expansion of $\left[\mathbf{I}-\mathcal{T}_{\boldsymbol{\mu}^v}\right]^{-1}\mathcal{T}_{\boldsymbol{\kappa}}$

$$\left[\mathbf{I}-\mathcal{T}_{\boldsymbol{\mu}^v}\right]^{-1}\mathcal{T}_{\boldsymbol{\kappa}} = \mathcal{T}_{\boldsymbol{\kappa}} + \mathcal{T}_{\boldsymbol{\mu}^v}\mathcal{T}_{\boldsymbol{\kappa}} + (\mathcal{T}_{\boldsymbol{\mu}^v})^2\mathcal{T}_{\boldsymbol{\kappa}} + (\mathcal{T}_{\boldsymbol{\mu}^v})^3\mathcal{T}_{\boldsymbol{\kappa}} + \dots. \tag{4.10}$$

Different from the logic of the standard Leontief inverse expansion, $(\mathbf{I}-\mathbf{A})^{-1}$, condition (4.10) incorporates two two novel ingredients of the model: (i) the attention choices are endogenous to both shock volatility and government policy; (ii) the attention linkages hinges on production networks due to general equilibrium feedbacks and strategic complementarity in information acquisition. These features not only provide an endogenous account of observed heterogeneity in sectoral browsing activities, they also point to a novel feedbacks between endogenous rigidities and optimal policy which we explore in the sequel.

Optimal monetary policy. In the first-order condition (4.2), the impact of endogenous attention channel also depends on the welfare exposure to attention variation, $\frac{\partial L}{\partial \boldsymbol{\mu}}$. The following lemma provides the necessary formula.

Lemma 4.3. *The welfare exposure to sectoral attention changes is*

$$\frac{\partial L}{\partial (\boldsymbol{\mu}^v)'} = \boldsymbol{r}^o + \boldsymbol{r}^c + \boldsymbol{r}^d. \tag{4.11}$$

Denote $\mathbf{M} = (\mathbf{I}-\mathbf{A})^{-1}(\boldsymbol{\mu}-\mathbf{I})\,\mathbb{V}(\boldsymbol{mc}_t)$, *where* $\mathbb{V}(\boldsymbol{mc}_t) \geq 0$ *denotes the covariance matrix of the sectoral marginal costs. The* $1 \times N$ *vectors* $\boldsymbol{r}^o$ *,* $\boldsymbol{r}^c$*, and* $\boldsymbol{r}^d$ *are given by*

$$\begin{aligned}
\boldsymbol{r}^o &= \frac{1}{\gamma + 1/\eta}(\boldsymbol{\beta}'\mathbf{M}) \odot \left(\boldsymbol{\beta}'(\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}\right), \\
\boldsymbol{r}^c &= \boldsymbol{\lambda}'\left[\mathbf{M} \odot (\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}\right] - \boldsymbol{\lambda}'\left[(\mathbf{A}\mathbf{M}) \odot \left(\mathbf{A}(\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}\right)\right] - \boldsymbol{r}^o, \\
\boldsymbol{r}^d &= \frac{1}{2}\boldsymbol{\chi}'.
\end{aligned}$$

Here, the vector $\boldsymbol{r}^o$ represents the impact through the output gap volatility, $\boldsymbol{r}^c$ through the cross-sector pricing error dispersion, and $\boldsymbol{r}^d$ through the within-sector pricing error dispersion.

Now we are ready to characterize the optimal monetary policy rule. Due to that the endogenous attention is the fixed point of the best response $\mathcal{T}(\boldsymbol{\mu}, \boldsymbol{\kappa})$, the solution to the first-order condition (4.2) is highly non-linear. Despite of this difficulty, the optimal policy still admit a compact price stabilization representation.

Proposition 4.3. *The optimal monetary policy can be implemented by a price-stabilization policy of the form*

$$\sum_{i=1}^{N} \varphi_i p_{it} = 0, \quad \textit{with} \quad \varphi_i = \varphi_i^x + \varphi_i^e, \tag{4.12}$$

where φ_i^e represents the weight due to endogenous attention given by

$$\varphi_i^e = \left\{ 2 \left[\sum_{j=1}^{N} \left(r_j^o + r_j^c + r_j^d \right) [\mathbf{I} - \mathcal{T}_{\boldsymbol{\mu}^v}]_{ij}^{-1} \right] - \left(\lambda_i \theta_i \mathbb{V}\left(mc_{it} \right) - \chi_i \right) \right\} \frac{\rho_i}{\mu_i} \frac{1 - \mu_i}{\mathbb{V}(mc_{it})}, \tag{4.13}$$

and φ_i^x given by (4.4) represents the weight when holding the attention constant.

The endogenous component of the policy rule (4.13) highlights the role of expectation management and network structure. The weight on a sector φ_i^e depends on the welfare exposures to price flexibilities ($\boldsymbol{r}$) in all other sectors with their importance adjusted by elements in the matrix $[\mathbf{I} - \mathcal{T}_{\boldsymbol{\mu}^v}]^{-1}$, showing that the policemaker internalizes strategic complementarity in firms' attention choices in the network setting. Meanwhile, the weight also inversely depends on the volatility of sectoral marginal costs, $\mathbb{V}(\mathrm{mc}_{it})$, which happens as the price flexiblities is no longer an exogenous parameter but depends on the equilibrium dynamics. Lastly, the term $(\lambda_i \theta_i \mathbb{V}\left(\mathrm{mc}_{it}\right) - \chi_i)$ corrects the discrepancy between elastic and fixed information capacity.

Meanwhile, what hidden from the this formula is the endogenous price flexibilities $\boldsymbol{\mu}$ induced by the optimal policy. Given the policy rule $\boldsymbol{\varphi}$, the corresponding nominal wage function $w_t = \boldsymbol{\kappa} \boldsymbol{z}_t$ can be constructed as

$$\boldsymbol{\kappa} \propto \boldsymbol{\varphi}(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}, \quad \text{subject to} \quad \boldsymbol{\kappa}\boldsymbol{\alpha} = 1. \tag{4.14}$$

Recall that in the best response function, $\mu_i = \mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa})$, firms' attention allocation hinges on the perceived wage function. While under the optimal monetary policy, the wage function is also determined by the attention allocation as in (4.14). The feedback between different policy rules and firms' active information acquisition decisions will lead to different equilibrium price rigidities, an aspect that is missing in an environment with exogenous information or fixed information capacity. Next, we will explore the properties of the implied price

4.3 Policy Induced Rigidity: Understanding the Mechanism

In this subsection, we study implications of optimal policy design on firms' attention and implied nominal rigidity. We proceed by exploring three exercises where each time the policymaker minimizes only one of the three components in the social loss function (4.1). We will show that relative to the exogenous-information or fixed information capacity economy, the optimal policy with elastic attention induce actually more nominal rigidity.

Minimizing output gap volatility. We start with the case in which the central bank only attempts to resolve welfare losses associated with the inefficient output-gap volatility. Surprisingly, the optimal policy rule in this case takes a particularly simple form as presented in the following proposition.

Proposition 4.4. *When minimizing only the output gap volatility, the optimal policy weight becomes*

$$\varphi_i = \frac{(1-\rho_0)}{(\gamma+1/\eta)}\lambda_i\left(\frac{1}{\mu_i}-1\right). \tag{4.15}$$

if at the optimum, the output-gap volatility is completely stabilized , $\mathbb{V}\left(c_t - c_t^*\right) = 0$.

Notice that when only minimizing the output gap, if the output-gap volatility is completely stabilized , $\mathbb{V}\left(c_t - c_t^*\right) = 0$, the optimal policy rule with exogenous information and elastic attention coincides with each other.[22] The price stabilization rule requires to eliminate the aggregate pricing error, $\sum_j \beta_j e_{jt} = 0$, which implies that the welfare exposure vector $\boldsymbol{r}^o = \mathbf{0}$ and the endogenous component $\varphi_i^e = 0$. This envelop-type result significantly simplifies the exposition.[23] However, this result by no means suggests that the equilibrium allocation is the same under elastic attention versus alternative information structures. Note that under elastic attention, μ_i in condition (4.15) still needs to be determined endogenously, while it is determinedly in an exogenous way with either fixed information capacity or exogenous information.

Particularly, suppose we start with the exogenous-information case with a vector of $\{\mu_i\}$. Formula (4.15) says that it is optimal to put more weight to stabilize sectors with smaller μ_i or more rigidity. As a result, the volatility of p_{it} in the more rigid sector tends to be smaller. If sector i uses their own sector's products intensively as intermediate inputs (which is the case empirically), the volatility of the marginal cost mc_{it} will be lower as well. Now let us switch to our baseline environment with elastic attention: with a less volatile marginal cost, sector i tends to pay less attention and adopts an even smaller μ_i, which further encourages policymaker to put more weight on sector i according to formula (4.15). This feedback effects between the optimal policy and firms' information acquisition best response creates a reinforcing channel that endogenously amplifies nominal rigidities.

Minimizing within-sector dispersion. The same logic also applies if the policymaker only tries to minimize the within-sector dispersion component in the welfare loss function. With this single objective, the sum of the endogenous and exogenous component in the price stabilizing rule also leads to a much simpler formula.

Proposition 4.5. *When minimizing only the within-sector dispersion, the optimal policy weight becomes*

$$\varphi_i = \left(\sum_{j=1}^{N}\chi_j\mathcal{R}_{ji}\right)\frac{\rho_i}{\mathbb{V}(mc_{it})}\left(\frac{1}{\mu_i}-1\right) \tag{4.16}$$

where $\mathcal{R}_{ji}$ is elements of the matrix $\left[\mathbf{I}-\mathcal{T}_{\boldsymbol{\mu}^v}\right]^{-1}$.

[22]In our quantitative analysis, we find the optimal OG policy indeed satisfies this condition under model calibration.

[23]Note that when the planner is simultaneously reducing the cross-sector or within-sector pricing error dispersion, $\boldsymbol{r}^o = \mathbf{0}$ ceases to be the case.

Fixing the information acquisition complementarity matrix $\mathcal{R}$, a lower μ_i or less volatile marginal cost implies translates into a larger weight in sector i. More stabilized price in sector i further reduces μ_i and $\mathbb{V}(\text{mc}_{it})$ through endogenous attention choice. As a result, the policy-attention feedback points to the same implication on the implied nominal rigidities.[24]

Also recall from Proposition 4.1, with exogenous information or fixed information capacity, the policy weight is $\varphi_i = \lambda_i \theta_i \rho_i (1 - \mu_i)$. This policy rule also implies a higher weight on more rigid sectors. Differently, there is no further modification of μ_i resulting from a higher φ_i.

Minimizing cross-sector dispersion. Lastly, if the policymaker only aims to eliminate the cross-sector dispersion, the only way is to encourage all firms to pay full attention, ($\mu_i = 1$), so that the sectoral pricing errors vanish. One way to achieve this goal is to set the monetary policy in a way such that the wage function exhibit extremely large response to all shocks, which results in extremely large volatility of sectoral marginal costs. Consequently, there is no longer an interior solution to the optimal policy problem, and the price stabilization rule fails to hold.

It is useful to note that though the cross-sector dispersion term alone will lead to a corner solution, when simultaneously considering all the three terms in objective (4.1) permits an interior solution. As we will show in the next section, the first two terms are dominating forces quantitatively in shaping the optimal monetary policy and most intuition can be obtained by inspecting the two special cases characterized by Proposition 4.4 and 4.5.

4.4 An Illustrating Example

To further illustrate the interaction between endogenous rigidities and optimal monetary policy, we consider an economy where all sectors are identical except for the volatility of productivity shocks. For simplicity, we abstract from the use of intermediate goods and assume that labor is the only production input. As a result, the nominal marginal cost in sector i is

$$\text{mc}_{it} = -z_{it} + w_t,$$

and firms' attention is given by $\mu_i = 1 - \frac{\chi}{\theta \lambda \mathbb{V}(\text{mc}_{it})}$. If the monetary policy implies a wage rule $\boldsymbol{\kappa}$ that is independent of firms' attention choice, then the sectoral rigidity dispersion will be pinned down by the volatility of productivity shocks and any strategic interactions among firms will be muted. One example of such symmetric and exogenous policy is that $\kappa_i = \frac{1}{N}$. However, this is not the case under the optimal monetary policy.

For simplicity, we focus on the special case where the policymaker only aims to minimize the aggregate output gap. Applying Proposition 4.4, the weight in the price stabilization rule is proportional to the inverse of sectoral attention, $\varphi_i \propto \frac{1}{\mu_i} - 1$. By condition (4.14), this monetary policy rule implies

[24] In our quantitative exercise in Section 5, we verify this intuition in our calibrated model.

the following nominal wage function that is endogenous to the sectoral rigidities, $\kappa_i = \frac{1-\mu_i}{\sum_j(1-\mu_j)}$.[25] Therefore, combining the best response of firms' attention choices and the policy minimizing the output gap leads to the following policy-induced fixed-point problem

$$\mu_i = 1 - \frac{\chi}{\theta\lambda}\frac{1}{(1-\kappa_i)^2\sigma_i^2 + \sum_{j\neq i}\kappa_j^2\sigma_j^2}, \tag{4.17}$$

$$\kappa_i = \frac{1-\mu_i}{\sum_j(1-\mu_j)}. \tag{4.18}$$

First assume momentarily that the wage function is symmetric across sectors, $\kappa_i = \frac{1}{N}$. When a sector i is with a relatively low productivity volatility (small σ_i^2), the volatility of their marginal cost is also relatively low, which implies firms in this sector have less incentive to acquire information and therefore choose a smaller μ_i according to (4.17). Next, we allow the wage function to respond to sectoral rigidities. According to (4.18), a smaller μ_i due to lower σ_i^2 now implies more policy emphasis or relatively large φ_i and κ_i, which in turn further reduces the volatility of sector i's marginal costs and incentivizes firms to pay even less attention. On the other hand, the sectors with volatility productivity shocks will end up paying even more attention with smaller φ_i and κ_i. The feedbacks between firms optimal attention allocation and the monetary policy reinforce each other and amplify the dispersion of sectoral rigidities.

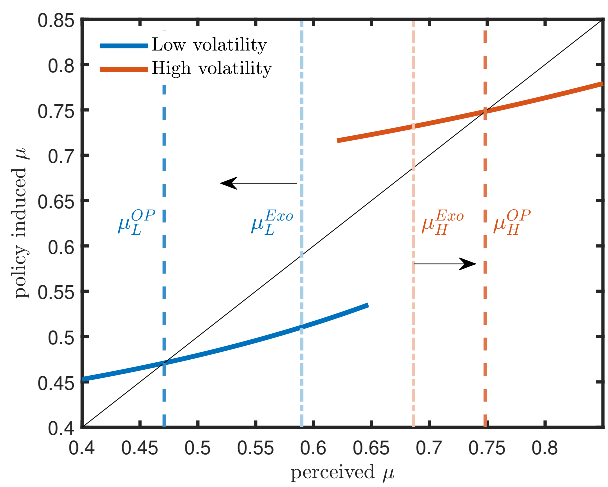


Figure 4: Optimal Policy Induced Rigidity Dispersion

To visualize the aforementioned mechanism, we divide the sectors into two groups: the volatility of productivity shocks in the first m sectors is lower than that of remaining m sectors, $\sigma_L^2 < \sigma_H^2$. With the exogenous wage function $\kappa_i = \frac{1}{N}$, the rigidities are pinned down by the shock volatility in their own sector. The broken blue line in Figure 4 represents the level of attention in the low volatility sectors (μ_L^{Exo}), while the red line represents the level of attention in the high volatility sectors (μ_H^{Exo}). When switching to the output gap minimizing policy, the dashed blue and red lines represent the

[25]This is obtained by using condition (4.14) with $\mathbf{A} = 0$ and $\alpha_i = \alpha = 1$ in our example.

new equilibrium level of attention allocation (μ_L^{OP} and μ_H^{OP}). Note that $\mu_L^{OP} < \mu_L^{Exp}$ and $\mu_H^{OP} > \mu_H^{Exp}$, which underscores the additional dispersion of sectoral rigidities with optimal monetary policy.

Furthermore, the solid blue and red lines illustrate the working of the fixed-point system (4.17) and (4.18). For example, the solid blue line shows that when holding μ_H at its equilibrium level μ_H^{OP}, a perceived μ_L in the monetary policy function (4.18) maps to the actual attention choice μ_L in (4.17). The fixed-point is then located at the cross with the 45 degree line. Similarly, the red solid line represents the mapping for high volatility sectors holding $\mu_L = \mu_L^{OP}$. In either case, the fixed-point mapping pushes the endogenous rigidity further away from each other.

5. Quantification

In this section, we provide a quantitative account of the theory. We show that the calibrated model is capable of replicating salient patterns of attention allocation as in firms' browsing activities. The model also predicts a positive correlation between sectoral price flexibility and sectoral shock volatility, which helps distinguish with alternative information structures. As anticipated in the previous subsection, we compare the optimal policy rule with that under exogenous information and quantify the effects of policy induced dispersion of endogenous price rigidities.

5.1 Calibration

The model is calibrated at a quarterly frequency. The parameters calibrated externally are based on conventional values in the literature. The elasticity of intertemporal substitution γ is set to 1. The Frisch elasticity η is calibrated to be 2. To achieve an average markup of 20%, the elasticity of substitution between goods within a sector θ is chosen to be 6. There are also a set of parameters that can be obtained directly from data. The elements in the production network matrix a_{ij} and the final goods shares β_i are computed base on the input-output table from BEA. The volatilities of sector-level productivity shocks are calibrated using the BEA/BLS Integrated Industry-Level Production Account (ILPA).

The calibration of the information acquisition costs χ_i is more unique in our setting. There are two challenges involving the choice of this set of parameters. First, most variables in the model such as price flexibilities are endogenous to policy. Therefore, the internal calibration requires us to take a stand on the monetary policy rule when bring the model to the data. As a baseline, we assume that the current monetary authority employs a CPI price stabilization rule, that is,

$$\sum_i \varphi_i p_{it} = 0, \quad \text{with} \quad \varphi_i = \frac{1}{\beta_i}.$$

We also explore alternative monetary policy rules, and the main results are robust to these alternatives.

Second, there are a large number of parameters to be determined as χ_i is sector specific. Motivated by condition (3.6), we allow χ_i to depend on λ_i and σ_i in a flexible way and impose the following parsimonious functional form

$$\log \chi_i = \delta_0 + \delta_1 \log \lambda_i + \delta_2 \log \sigma_i.$$

We calibrate the parameters δ_0, δ_1 and δ_2 to match the distribution of *adjusted* forecast errors of earnings per share (EPS) at the sector level, which are directly related to the informational frictions.[26] We choose EPS as our target because it provides a significantly larger number of observations compared to other forecasted variables in the IBES, allowing us to aggregate the forecast error at the sector level with small standard errors. Particularly, we target the mean level of the forecast error, the 25 percetile and 75 perticle of the forecast error in the sectoral distribution. Table 6 lists the calibrated parameters.

Table 6: Calibrated Parameters

Param.	Value	Source	Related to
Exogenously determined parameters			
γ	1	—	income elasticity
η	2	—	Frisch elasticity
θ	6	20% markup	elasticity of substitution
a_{ij}		BEA	input-output matrix
σ_i		KLEMS	productivity shock volatility
Endogenously determined parameters			
δ_0	0.30	sectoral forecast error of EPS	information acquisition cost
δ_1	0.71	sectoral forecast error of EPS	information acquisition cost
δ_2	1.60	sectoral forecast error of EPS	information acquisition cost

As shown in Table 7, though not directly targeted, our model matches the distribution of the implied price-change frequency reasonably well. The model yields similar amount of sectoral price rigidities as the data at different percentiles.

5.2 Model v.s. Data: Attention, Volatility, and Price Flexibility

In this subsection, we compare various aspects of the calibrated model predictions with their data counterparts and illustrate the role of endogenous information acquisition.

[26] Specifically, we choose δ_0, δ_1 and δ_2 to minimize the following loss function,

$$\min_{\delta_0,\delta_1,\delta_2} \left(\mathrm{FE}^{eps,data}_{i=25\%} - \mathrm{FE}^{eps,model}_{i=25\%}\right)^2 + \left(\sum_{i=1}^{N} \frac{1}{N}\mathrm{FE}^{eps,data}_{i} - \sum_{i=1}^{N} \frac{1}{N}\mathrm{FE}^{eps,model}_{i}\right)^2 + \left(\mathrm{FE}^{eps,data}_{i=75\%} - \mathrm{FE}^{eps,model}_{i=75\%}\right)^2,$$

, where FE_i is the mean absolute forecast error of sector i. However, the volatility of EPS differs in the data and in the model. To ensure that the forecast error is comparable, we normalize it by the volatility of EPS, both in the data and in the model.

Table 7: Model Fit: Forecast Error and Price-Change Frequency

Moment	Forecast Error		Frequency	
	Data	Model	Data	Model
Moments targeted				
Mean	0.28	0.27		
25th percentile	0.20	0.22		
75 percentile	0.33	0.33		
Moments not targeted				
Mean			0.23	0.24
Standard deviation	0.12	0.08	0.13	0.11
10th percentile	0.13	0.18	0.12	0.12
25th percentile			0.14	0.16
50th percentile	0.26	0.27	0.19	0.23
75 percentile			0.29	0.30
90 percentile	0.48	0.36	0.36	0.39

Attention allocation. As documented in Section 2, a salient feature of firms' browsing activities is that they are positively associated with both the input-output linkages and sectoral shock volatilites. In our model, firms acquire signals about their nominal marginal costs, which are informativeness about the underlying sectoral shocks. To see to which extent our model can replicate the empirical regularity, we leverage Definition 2, where the attention allocated to a sector j from sector i, ω_{ij}, is measured by the reduction of posterior uncertainty of sector i's productivity shock relative to its prior. Consistent with the empirical specification (2.5), we run the following regression with model generated data

$$\omega_{ij} = \alpha_i + \beta_1 \text{ input\_share}_{ij} + \beta_2 \text{ sales\_share}_{ij} + \beta_3 \text{ sales\_volatility}_j + \varepsilon_{ij}.$$

Different from its empirical counter, the model is stationary and the independent variables are time invariant.

Table 8 displays the regression coefficients. The model yields a qualitatively similar pattern as that in the data (Table 4 column (5)): the percentage reduction of uncertainty about a sector is positively related with the sectoral sales volatility and bilateral trade linkages. This is much anticipated based on Proposition 3.7, which states that ω_{ij} is determined by the product of equilibrium exposure of sector i's price to sector j's shock and the shock volatility in sector j. This pattern underscores the importance of the endogeneity of firms' information acquisition choice. For example, with exogenous information such as La'O and Tahbaz-Salehi (2022), though firms in different sectors face different degrees of information frictions, the attention of a firm is allocated in a uniform way across sectors, that is, $\omega_{ij} = \omega_{ik}$.

Table 8: Attention Allocation: Sectoral-Level Regression

	Regression Coefficient		
	input share	sales share	volatility
Data	0.34***	0.33***	0.03***
	(0.07)	(0.06)	(0.01)
Model	1.28***	0.62***	0.07***
	(0.05)	(0.03)	(0.02)

Price flexibility and shock volatility. Another important implication of the endogenous information acquisition channel is that the implied price flexibilities are directly connected with the underlying shock volatilities. This feature is absent in models with Calvo-type frictions or exogenous information frictions, in which case the price flexibilities are exogenously determined and independent of fundamentals.

Figure 5: Frequency and Volatility

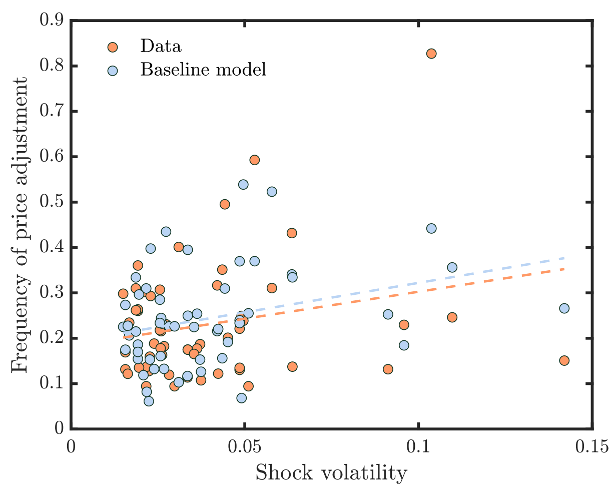
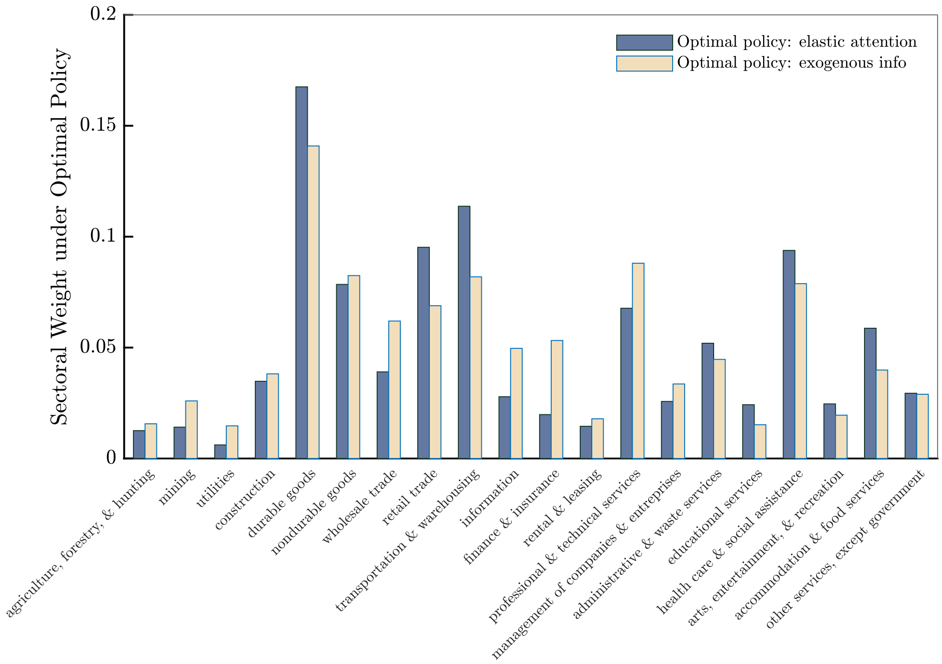


Figure 5 explores such connection in both the data and our calibrated model. Each dot in the figure represents a sector's own productivity shock volatility and its frequency of price adjustment. The red ones show that in the data, a higher shock volatility is associated with greater price flexibility. The blue ones display the same relationship in the model. Though our calibration predicts a higher frequency of price adjustment on average as discussed in previous subsection, it yields a positive relationship with a similar slope. The reason for this pattern is straightforward: condition (3.5) predicts that the price flexibility is increasing in the volatility of the marginal cost, and the own productivity shock is quantitatively important in driving the marginal cost.

Figure 6: Optimal Monetary Policy Rule



5.3 Allocation under Optimal Monetary Policy

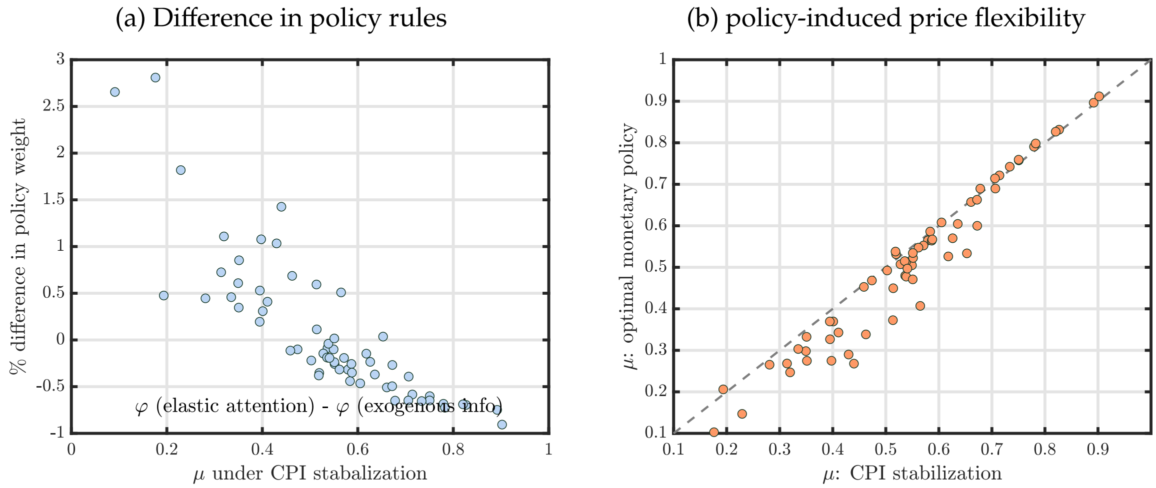
With the calibrated model, we now quantify the effects of endogenous information acquisition in shaping the optimal monetary policy and the policy-induced price flexibilities.

We start with the comparison of the monetary policy rule with elastic attention in our baseline model with that under exogenous information and fixed information capacity. To make these economies comparable with each other, we invoke Proposition 3.3 and set the maximum mutual information capacity $\{\delta_i\}$ and the variance of noise $\{\tau_i\}$ so that the implied price flexibilities $\{\mu_i\}$ are identical to that in our baseline calibration with elastic attention and CPI price stabilization. Due to Proposition 4.1, the optimal monetary policy rule with exogenous information coincides that with fixed information capacity, and we will only refer to the exogenous information case when making comparison with our baseline model.

Figure 6 displays the sectoral weights under the optimal price stabilization rule, both of which are normalized so that they sum to one. The beige bars represent the weights with exogenous information according to Proposition 4.1, where the informational friction is independent of the policy. The blue bars represent the weights in our baseline model with elastic attention according to Proposition 4.3, where the policy weights contains the additional component due to the additional expectation management motive. The two policy rules are significantly different, with an average percentage difference change equaling 24.3%, which suggests that the nature of the underlying informational frictions matter for the design of optimal policy rule.

The next question is what factor drives the differences between these two policy rules. To answer this question, Panel (a) of Figure 7 plots the percentage difference of the optimal sectoral weights between the exogenous information economy and the elastic attention economies against the initial price flexibilities under CPI stabilization. Anticipated from the discussion in last section, there is a sharp pattern that the difference is amplified when the initial level of price flexibility is low, and the heterogeneity in the price flexibilities account for more than 85% of the policy weight difference. With exogenous information, the more rigid sectors receive a higher weight. With elastic attention, the policymaker will put an additional emphasis on these sectors due to the feedback effects between the policy and the endogenous flexibilities, which verifies the intuition developed in Proposition 4.4 and 4.5 in the full calibrated model.

Figure 7: Optimal Policy Rule and Endogenous Price Flexibility



Panel (b) of Figure 7 further highlights the aforementioned endogenous response of price flexibilities. The sectors that are relatively more rigid (flexible) under the CPI stabilization policy rule becomes even more so under the optimal monetary policy rule. The dispersion of price rigidities is therefore amplified. Again, this result is consistent with the findings in the stylized example in Subsection 4.4.

6. Conclusion

This paper explores how firms acquire information in complex production networks. Utilizing novel data sets on firms' browsing activities; we establish an "attention network" in the economy and present three empirical facts that characterize this network. Motivated by the empirical evidence, we develop a theoretical framework featuring a production network, rationally inattentive firms, and monetary policy. In our framework, the input-output linkages, the volatilities of sectoral shocks, and the monetary policy endogenously determine sectoral nominal rigidities, information acquisition, and strategic interactions of sectoral attentions in general equilibrium. These features allow the model to rationalize our empirical findings, both qualitatively and quantitatively. In particular, our model parsimoniously captures the cross-sector distribution of forecast errors and nominal rigidity observed in the data. We also analyze the design of optimal monetary policy in this environment. When the central bank can manage firms' expectations (and therefore nominal rigidities), we find that the optimal policy implementation differs substantially from the model with exogenous information frictions: feedback between optimal policy and attention leads to endogenous dispersion in the cross-sector distribution of nominal rigidities.

Several research directions remain to be explored in the future. First, our analysis focuses on a static setting; it would be interesting to investigate how a dynamic environment affects firms' information acquisition and the design of optimal monetary policy in the context of production networks. In this case, persistent features of shocks and dynamic impacts of monetary policy will add an extra dimension of complexity to the model analysis. Second, our current results mainly speak about the nominal side of the economy. How endogenous information acquisition influences sectoral shock propagation and real economic allocations (e.g., Hulten's theorem) remains an open question. To address this question, adopting more general CES production technology that generalizes our Cobb-Douglas framework seems necessary. The underlying research directions go beyond the scope of this paper, and we leave it for future work.

References

Adam, K. (2007): "Optimal monetary policy with imperfect common knowledge," *Journal of Monetary Economics*, 54, 267–301.

Adams, J. (2019): "Macroeconomic Models with Incomplete Information and Endogenous Signals," Working Papers 001004, University of Florida, Department of Economics, `https://EconPapers.repec.org/RePEc:ufl:wpaper:001004`.

Afrouzi, H. (2023): "Strategic Inattention, Inflation Dynamics, and the Non-Neutrality of Money," Working Paper 31796, National Bureau of Economic Research, 10.3386/w31796.

Angeletos, G.-M., and Z. Huo (2021): "Myopia and Anchoring," *American Economic Review*, 111, 1166–1200, 10.1257/aer.20191436.

Angeletos, G.-M., L. Iovino, and J. La'O (2020): "Learning over the business cycle: Policy implications," *Journal of Economic Theory*, 190, 105115, https://doi.org/10.1016/j.jet.2020.105115.

Angeletos, G.-M., and K. Sastry (2019): "Inattentive Economies," Working Paper 26413, National Bureau of Economic Research, 10.3386/w26413.

Aoki, K. (2001): "Optimal monetary policy responses to relative-price changes," *Journal of Monetary Economics*, 48, 55–80.

Auclert, A., M. Rognlie, and L. Straub (2020): "Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model," *Working Paper*.

Bass, F. M. (1969): "A new product growth for model consumer durables," *Management science*, 15, 215–227.

Benigno, P. (2004): "Optimal monetary policy in a currency area," *Journal of International Economics*, 63, 293–320.

Berman, A., and R. J. Plemmons (1994): *Nonnegative Matrices in the Mathematical Sciences*, Classics in Applied Mathematics: Society for Industrial and Applied Mathematics, 10.1137/1.9781611971262.

Bui, H., Z. Huo, A. Levchenko, and N. Pandalai-Nayar (2024): "Noisy Global Value Chains," *Working Paper*.

Candia, B., O. Coibion, and Y. Gorodnichenko (2023): "The macroeconomic expectations of firms," in *Handbook of Economic Expectations*: Elsevier, 321–353.

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, 3603–3642.

Caplin, A., and M. Dean (2015): "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 105, 2183–2203, 10.1257/aer.20140117.

Caplin, A., M. Dean, and J. Leahy (2018): "Rational Inattention, Optimal Consideration Sets, and Stochastic Choice," *The Review of Economic Studies*, 86, 1061–1094, 10.1093/restud/rdy037.

——— (2022): "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy," *Journal of Political Economy*, 130, 1676–1715, 10.1086/719276.

Carvalho, C., J. W. Lee, and W. Y. Park (2021): "Sectoral Price Facts in a Sticky-Price Model," *American Economic Journal: Macroeconomics*, 13, 216–56, 10.1257/mac.20190205.

Carvalho, V. M., M. Nirei, Y. U. Saito, and A. Tahbaz-Salehi (2021): "Supply chain disruptions: Evidence from the great east japan earthquake," *The Quarterly Journal of Economics*, 136, 1255–1321.

Chahrour, R., K. Nimark, and S. Pitschner (2021): "Sectoral Media Focus and Aggregate Fluctuations," *American Economic Review*, 111, 3872–3922, 10.1257/aer.20191895.

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, 118–137.

Culot, G., M. Podrecca, G. Nassimbeni, G. Orzes, and M. Sartor (2023): "Using supply chain databases in academic research: A methodological critique," *Journal of Supply Chain Management*, 59, 3–25.

Diamond, D. W., and R. E. Verrecchia (1991): "Disclosure, liquidity, and the cost of capital," *The journal of Finance*, 46, 1325–1359.

Flynn, J. P., and K. Sastry (2019): "Attention Cycles," working paper.

——— (2023): "Attention cycles," *Available at SSRN 3592107*.

Francis, J., D. Nanda, and P. Olsson (2008): "Voluntary disclosure, earnings quality, and cost of capital," *Journal of accounting research*, 46, 53–99.

Ghassibe, M. (2021): "Monetary policy and production networks: an empirical investigation," *Journal of Monetary Economics*, 119, 21–39.

Guay, W., D. Samuels, and D. Taylor (2016): "Guiding through the fog: Financial statement complexity and voluntary disclosure," *Journal of Accounting and Economics*, 62, 234–269.

Hébert, B., and J. La'O (2023): "Information Acquisition, Efficiency, and Nonfundamental Volatility," *Journal of Political Economy*, 131, 2666–2723, 10.1086/724575.

Hellwig, C., and L. Veldkamp (2009): "Knowing What Others Know: Coordination Motives in Information Acquisition," *The Review of Economic Studies*, 76, 223–251, 10.1111/j.1467-937X.2008.00515.x.

Horn, R. A., and C. R. Johnson (2012): *Matrix analysis*: Cambridge University Press, 2nd edition.

Huang, K. X., and Z. Liu (2005): "Inflation targeting: What inflation rate to target?" *Journal of Monetary Economics*, 52, 1435–1462.

Hutson, V., J. S. Pym, and M. J. Cloud (2005): *Applications of Functional Analysis and Operator Theory*, Mathematics in Science and Engineering: Elsevier.

Jamilov, R., A. Kohlhas, O. Talavera, and M. Zhang (2024): "Granular Sentiments," Discussion Papers 2414, Centre for Macroeconomics (CFM), `https://ideas.repec.org/p/cfm/wpaper/2414.html`.

Kellogg, R. B. (1976): "Uniqueness in the Schauder Fixed Point Theorem," *Proceedings of the American Mathematical Society*, 60, 207–210, 10.2307/2041143.

La'O, J., and A. Tahbaz-Salehi (2022): "Optimal Monetary Policy in Production Networks," *Econometrica*, 90, 1295–1336, 10.3982/ECTA18627.

Li, M., and H.-M. Wu (2016): "Optimal Monetary Policy with Asymmetric Shocks and Rational Inattention," *mimeo*.

Luo, Y., J. Nie, G. Wang, and E. R. Young (2017): "Rational inattention and the dynamics of consumption and wealth in general equilibrium," *Journal of Economic Theory*, 172, 55–87, https://doi.org/10.1016/j.jet.2017.08.005.

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

Maćkowiak, B., and M. Wiederholt (2009): "Optimal sticky prices under rational inattention," *American Economic Review*, 99, 769–803.

——— (2015): "Business cycle dynamics under rational inattention," *The Review of Economic Studies*, 82, 1502–1532.

Magnus, J. R., and H. Neudecker (2007): *Matrix differential calculus with applications in statistics and econometrics*, Wiley Series in Probability and Statistics: Wiley, 3rd edition.

Matějka, F., and A. McKay (2015): "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105, 272–98.

Miao, J., J. Wu, and E. R. Young (2022): "Multivariate rational inattention," *Econometrica*, 90, 907–945.

Milgrom, P. R. (1981): "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 380–391.

Ou, S., P. Yin, D. Zhang, and R. Zhang (2024): "Nominal Rigidities, Rational Inattention, and the Optimal Price Index Stabilization Policy," working paper.

Paciello, L., and M. Wiederholt (2013): "Exogenous Information, Endogenous Information, and Optimal Monetary Policy," *The Review of Economic Studies*, 81, 356–388, 10.1093/restud/rdt024.

Pasten, E., R. Schoenle, and M. Weber (2020): "The propagation of monetary policy shocks in a heterogeneous production economy," *Journal of Monetary Economics*, 116, 1–22.

——— (2024): "Sectoral Heterogeneity in Nominal Price Rigidity and the Origin of Aggregate Fluctuations," *American Economic Journal: Macroeconomics*, 16, 318–352.

Pastén, E., R. Schoenle, and M. Weber (2024): "Sectoral Heterogeneity in Nominal Price Rigidity and the Origin of Aggregate Fluctuations," *American Economic Journal: Macroeconomics*, 16, 318–52, 10.1257/mac.20210460.

Plemmons, R. J. (1977): "*M*-matrix characterizations.I—nonsingular *M*-matrices," *Linear Algebra and its Applications*, 18, 175–188, 10.1016/0024-3795(77)90073-8.

Qaiser, S., and R. Ali (2018): "Text mining: use of TF-IDF to examine the relevance of words to documents," *International Journal of Computer Applications*, 181, 25–29.

Rogers, J. L., D. J. Skinner, and S. L. Zechman (2017): "Run EDGAR run: SEC dissemination in a High-Frequency world," *Journal of Accounting Research*, 55, 459–505.

Rubbo, E. (2023): "Networks, Phillips curves, and monetary policy," *Econometrica*, 91, 1417–1455.

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

——— (2010): "Rational inattention and Monetary economics," *Handbook of Monetary Economics*.

Smith, H. L., and C. A. Stuart (1980): "A uniqueness theorem for fixed points," *Proceedings of the American Mathematical Society*, 79, 237–240.

Song, W., S. Stern et al. (2022): "Firm Inattention and the Efficacy of Monetary Policy: A Text-Based Approach,"Technical report, Bank of Canada.

Talman, L. A. (1978): "A Note on Kellogg's Uniqueness Theorem for Fixed Points," *Proceedings of the American Mathematical Society*, 69, 248–250, 10.2307/2042606.

Tanaka, T., K.-K. K. Kim, P. A. Parrilo, and S. K. Mitter (2017): "Semidefinite Programming Approach to Gaussian Sequential Rate-Distortion Trade-offs," *IEEE Transactions on Automatic Control*, 62, 1896–1910.

Vetter, W. J. (1973): "Matrix Calculus Operations and Taylor Expansions," *SIAM Review*, 15, 352–369, 10.1137/1015034.

Online Appendix

Inattentive Networks: Evidence and Theory

Yue Fang   Zhen Huo   Minghao Li   Jieran Wu

A. Data Appendix

A.1 The EDGAR Browsing Data

A.1.1 Constructing the Browsing Data

The EDGAR log files contain the full viewing history of the filed documents published on EDGAR, including the viewers' IP address, the timestamp of the request, and the page requested. These log files were made available online by the SEC at https://www.sec.gov/data/edgar-log-file-data-set. The SEC provides the first three octets of the viewers' IP address and anonymizes the fourth octet with a 3-character string that preserves the last octet's uniqueness without revealing the IP's full identity.[27] Following the method proposed by Chen et al. (2020), we decipher the last octet and reveal the full IP address. We further map each viewer's uncovered IP address to their true identity using services provided by ip-info.io, a leading IP information provider. The final sample contains a total of 713,157,510 unique IP addresses of 7,622 public companies in our sample period, representing 52.4% of all browsings of disclosed files recorded in EDGAR. The summary statistics of the baseline sample are presented in Table 1 in the main text.

A.1.2 The Surge of EDGAR usage

The year 2009 marks a noticeable milestone in terms of EDGAR adoption. Graph A.2 shows the Google Trend for the keyword "SEC EDGAR search", directly linking to EDGAR's search page, which is the primary point of entry for human users. The number of searches surged a few months prior to 2009, followed by a stable level of searches thereafter.

Despite the general rule of new product adoption which predicts a Bass Diffusion Model growth curve (Rogers et al. (2017); Bass (1969)), the information content of disclosures may also contribute to the increasing usage of the EDGAR platform. The reason is twofold. First, due to the changing regulatory landscape, SEC requires the disclosure of more and more content over time.[28] Figure A.1 illustrates the average lengths of 10-Ks and 10-Qs, two primary types of public disclosures. It can be seen that the lengths are steadily increasing, suggesting that the information in disclosures gets richer over the years. Due to the benefits of reducing information asymmetry and the associated benefits, such as lowering the cost of capital(Francis et al. (2008)), mitigating the adverse selection problems (Milgrom (1981); Diamond and Verrecchia (1991)), and offsetting the negative effects of complex financial statements on the information environment (Guay et al. (2016)), firms also have incentives to disclose more information voluntarily.

[27] An octet is an integer in the range [0, 256). For example, an IP address provided by the SEC is "10.191.131.ace", where the fourth octet "ace" masks the last three digits of the IP address.

[28] For instance, SEC started mandating firms to disclose unresolved SEC staff comments (2005), the effectiveness of internal control over financial reporting (2007), risk factors (Item 1A, 2006), mining safety (Item 4, 2011), oil and gas reserves in possession of energy companies, hedging policies (2018), and payment to US federal and foreign governments by natural resource extractors (2020).

Figure A.1: Average Length of 10-K and 10-Q Documents

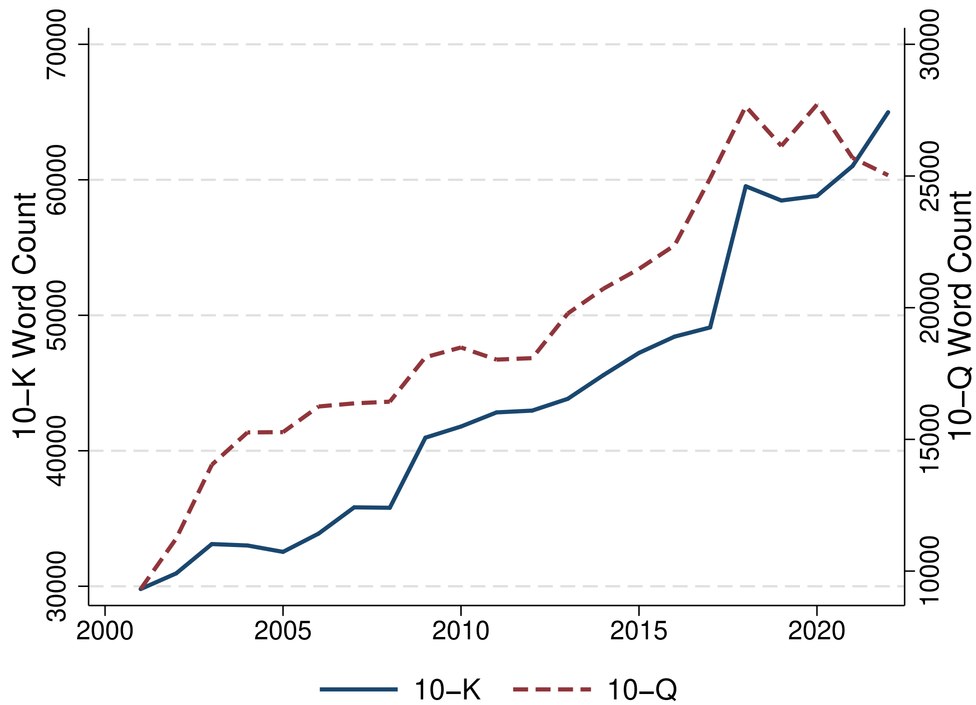
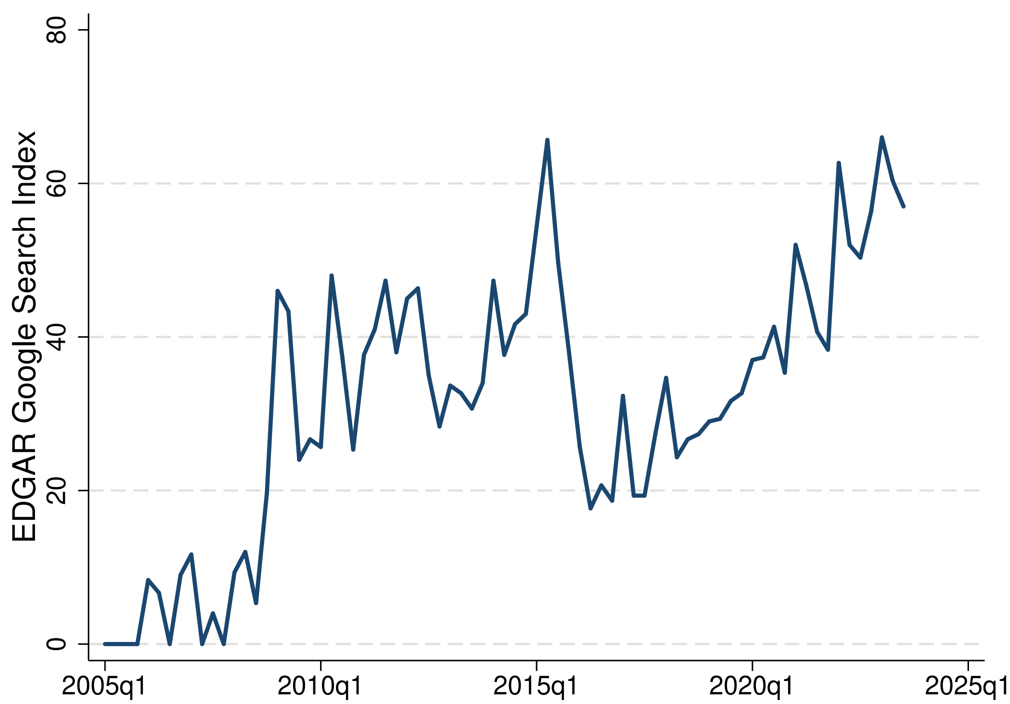


Figure A.2: EDGAR Google Search Index



A.2 Constructing Text-Based Attention Measure

10-K filings. The SEC requires public companies to file Form 10-K annually. The 10-K filing is a detailed document that provides a comprehensive overview of a company's financial performance, operations, and risks, among other important information. 10-K goes beyond regular annual reports and provides a more in-depth backward-looking and forward-looking analysis of a company.

Methodology. We also propose another measure of attention based on a text-based approach similar to that of Flynn and Sastry (2023) and Song et al. (2022). The idea is to choose a dictionary of "signal words" unique to each industry and calculate the frequency counts in a company's 10-K disclosures. For example, the word "uranium" is most likely to be mentioned when a firm is discussing the "Mining (except Oil and Gas)" industry (NAICS code 212). If a company repeatedly mentions "uranium" in its corporate disclosures (high word frequency), it likely devotes greater managerial attention to NAICS 212.

We select the keyword list for each industry i (defined by three-digit NAICS codes) following the steps below:

(1) Prepare corpus texts pertaining to each industry from two sources. First, we extract company descriptions for all public US companies from Osiris, a Bureau van Dijk data product that contains comprehensive information on public and private companies across the globe. In Osiris, each company has a short description of its history and an overview of its business and products. By matching these descriptions with their industry codes, they jointly portray the essence of the industries that they belong to. The second corpus is the descriptive texts of NAICS codes, maintained by the NAICS Association, LLC, a company specializing in NAICS code services.

(2) Pre-processing. Two python packages, Spacy and NLTK (Natural Language Toolkit) are prerequisites. We execute two procedures before running the main program. First, we use NLTK's embedded stop word list to remove common words like "the", "an", and "at" that lack meaning. Second, we run entity_recognization.py to leverage Spacy's Named Entity Recognition (NER) algorithm and the large English language model en_core_web_lg to identify and exclude geographic and geopolitical entities. This prevents some geographically-concentrated industries from capturing words like "Texas" and "China".

(3) After pre-processing, we tokenize the texts using NLTK's pre-trained language model and select all nouns (NN), adjectives (JJ), and verbs (VB) - the three parts of speech that form sentences. We first extract the 100 most common words across all industries (subjectively chosen but works well) and remove them, as they are unlikely to be unique to a specific industry.

(4) We employ the TF-IDF (Term Frequency-Inverse Document Frequency) algorithm, a method that is intended to reflect how relevant a word is to a document in a corpus (Qaiser and Ali, 2018), to select a dictionary of signal words Θ_j of each NAICS3 industry j. The intuition behind TF-IDF is that a term that appears frequently in a document but rarely in the rest of the corpus is more likely to be representative of the topic of the document. We vectorize the corpus documents, calculate the tf-idf score of every qualified word, and select the top 15 words with the highest tf-idf scores into the industry j's dictionary Θ_j.

(5) Using the dictionary generated in (4), We search for the word frequency of words in Θ_j for every j in firms' 10-K filings. We restrict the search to Item 1A (Risk Factors) and Item 7 (Management Discussion and Analysis), as companies are most likely to discuss an industry about their risk or strategic concerns in these two parts. Finally,

we aggregate the total number of occurrences of each word in industry j's dictionary Θ_j, indexed as k. Firm i's attention to industry j in year t is calculated as:

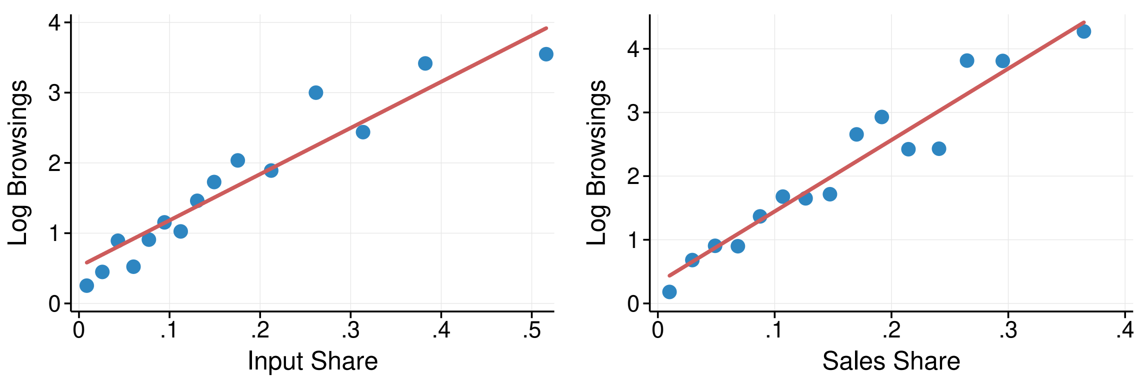
$$Attention_{i,j,t} = \sum_{k \in \Theta_j} WordFreq_{i,k,t} \tag{A.1}$$

B. Empirical Figures and Tables

Table B.1: Industry-level Summary Statistics

	71 Sectors			405 Sectors		
	Mean	Median	S.D.	Mean	Median	S.D.
Input Share	0.014	0.0025	0.040	0.0025	0.000029	0.014
Sales Share	0.014	0.0030	0.046	0.0025	0.000075	0.019
Inflation Volatility	0.10	0.053	0.15	0.091	0.049	0.12
Sales Growth Volatility	0.090	0.080	0.058	0.12	0.09	0.13
TFP Volatility	0.031	0.024	0.030			

Figure B.1: Browsing Intensity and Input-Output Linkage



Note: The left panel shows the binscatter plot between sector i's log browsings on an upstream sector j and j's share as a supplier of i. The right panel shows the binscatter plot between sector i's log browsings on a downstream sector j and j's share as a customer of i. Both plots remove the industry-year fixed effect.

Table B.2: Filed Documents Being Viewed Most

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

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

Table B.3: Browsing Intensity and Input-Output Linkages (Sales and TFP Volatility)

	Browsing Intensity					
	71 Sectors				405 Sectors	
	(1)	(2)	(3)	(4)	(5)	(6)
Input Share		4.28***		5.05***		11.0***
		(0.85)		(0.94)		(1.03)
Sales Share		5.88***		5.18***		9.35***
		(1.06)		(0.79)		(0.83)
Inflation Volatility	0.54***	0.52***			0.52***	0.55***
	(0.16)	(0.14)			(0.064)	(0.061)
TFP Volatility			5.80***	5.66***		
			(0.64)	(0.56)		
Industry Controls	✓	✓	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.79	0.81	0.79	0.80	0.67	0.69
No. Observations	23098	23098	26216	26216	311061	309238

Note: This table shows how the browsing intensity of an industry depends on its input share from the upstream sectors, the sales share from the downstream sectors, and how volatile the other sectors' sales or TFP are. Robust standard errors are in paraentheses and clustered by industry. *Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.4: Browsing Intensity and Forecast Accuracy

	Forecast Error			
	Earnings Per Share		Sales	
	(1)	(2)	(3)	(4)
Browsing Intensity (30 days)	-0.020**		-0.0011***	
	(0.0073)		(0.00023)	
Browsing Intensity (180 days)		-0.042***		-0.0016**
		(0.0090)		(0.00056)
Controls	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Adjusted R^2	0.25	0.25	0.42	0.42
No. Observations	6722	6722	7607	7607

Note: Standard errors are in parentheses and are clustered at industry level. This table shows how firms' forecast accuracy is assocatiated with firms' browsing intensity in the past 30 days or 180 days before the forecast day. .*Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure B.2: Distribution of Browsing Activity (per employment)

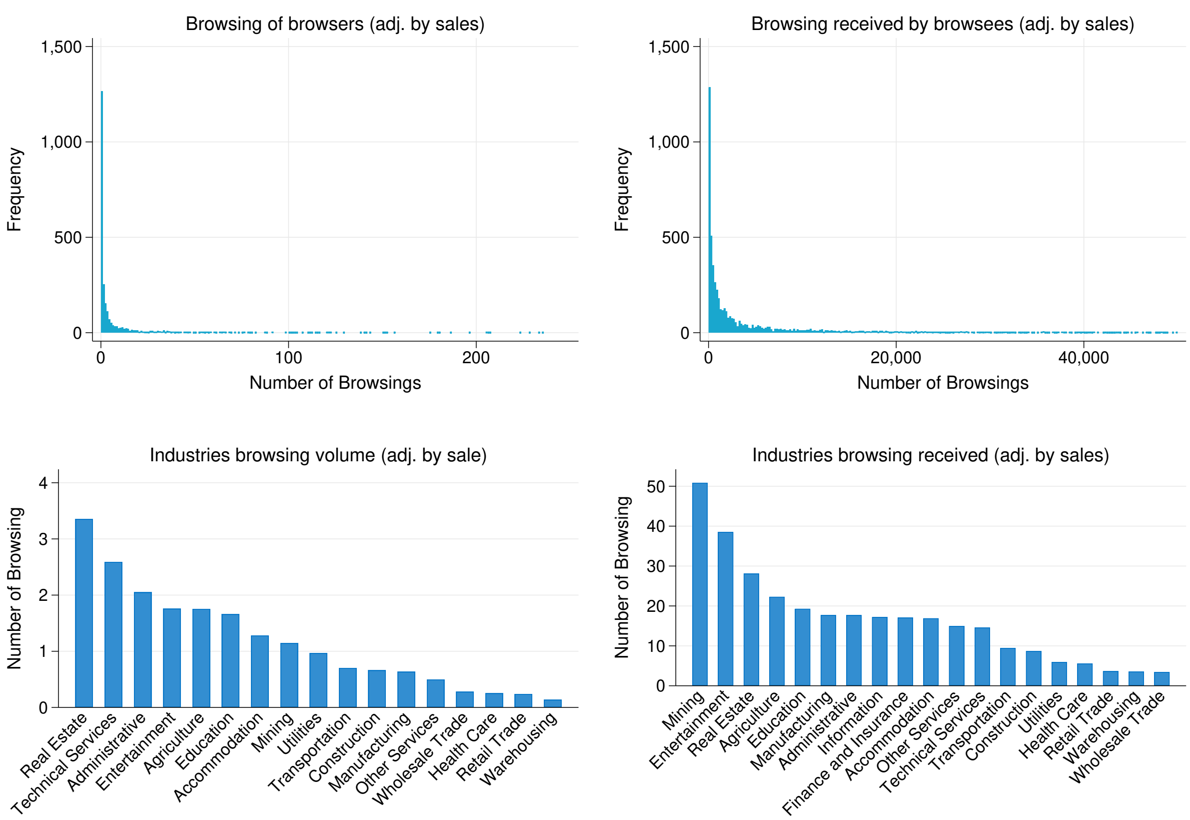


Figure B.3: Distribution of Browsing Activity (per sales)

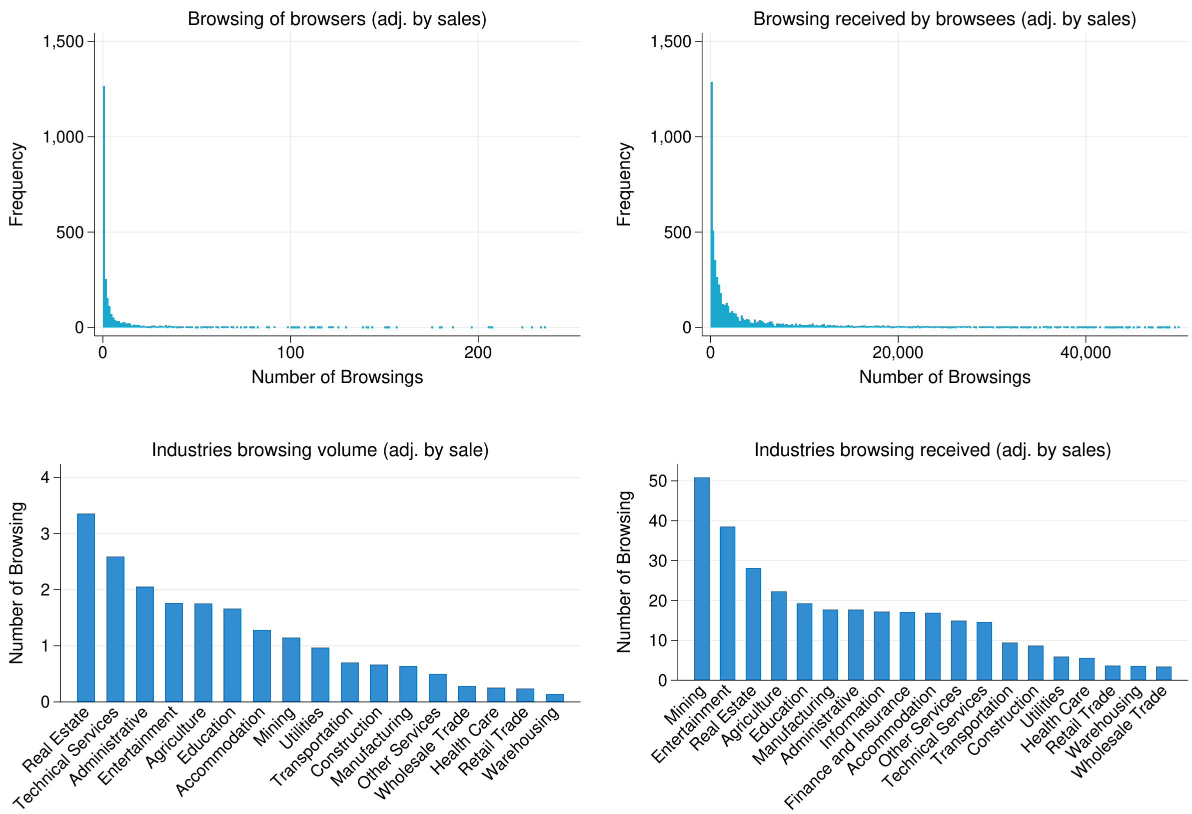


Figure B.4: Browsing Volume, Employment and Sales

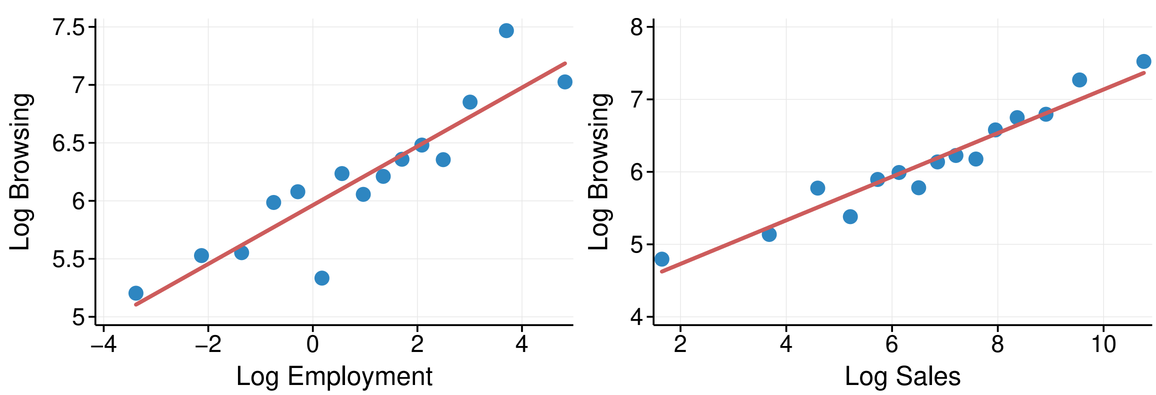
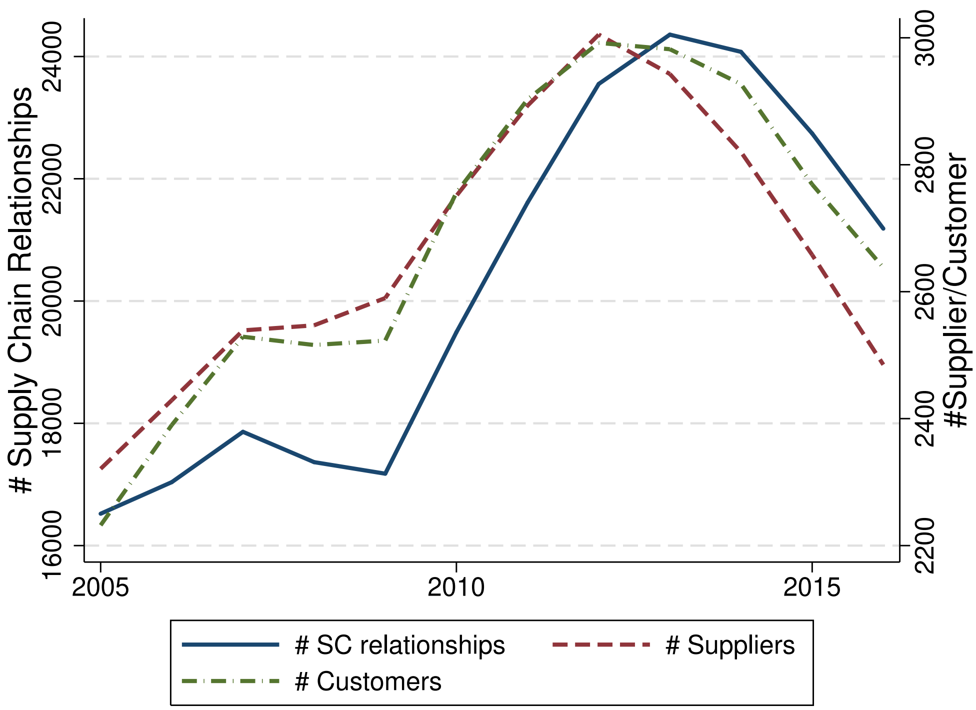
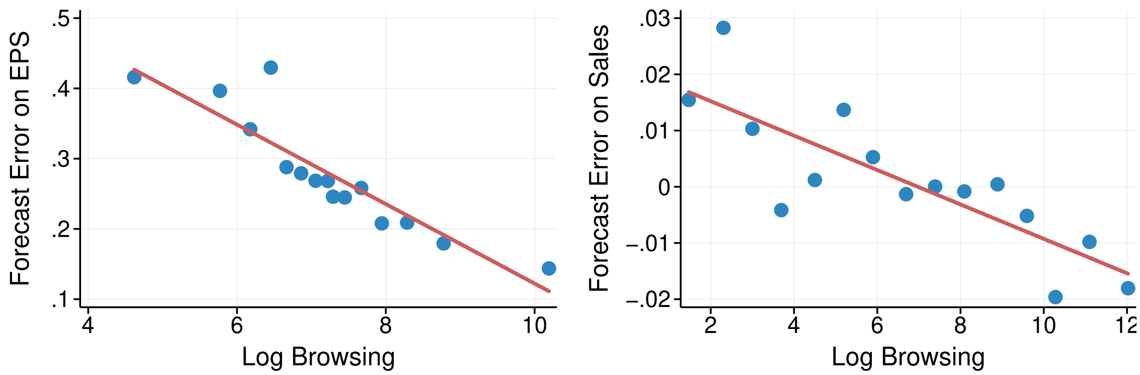


Figure B.5: Supply Chain Relationships Over Time



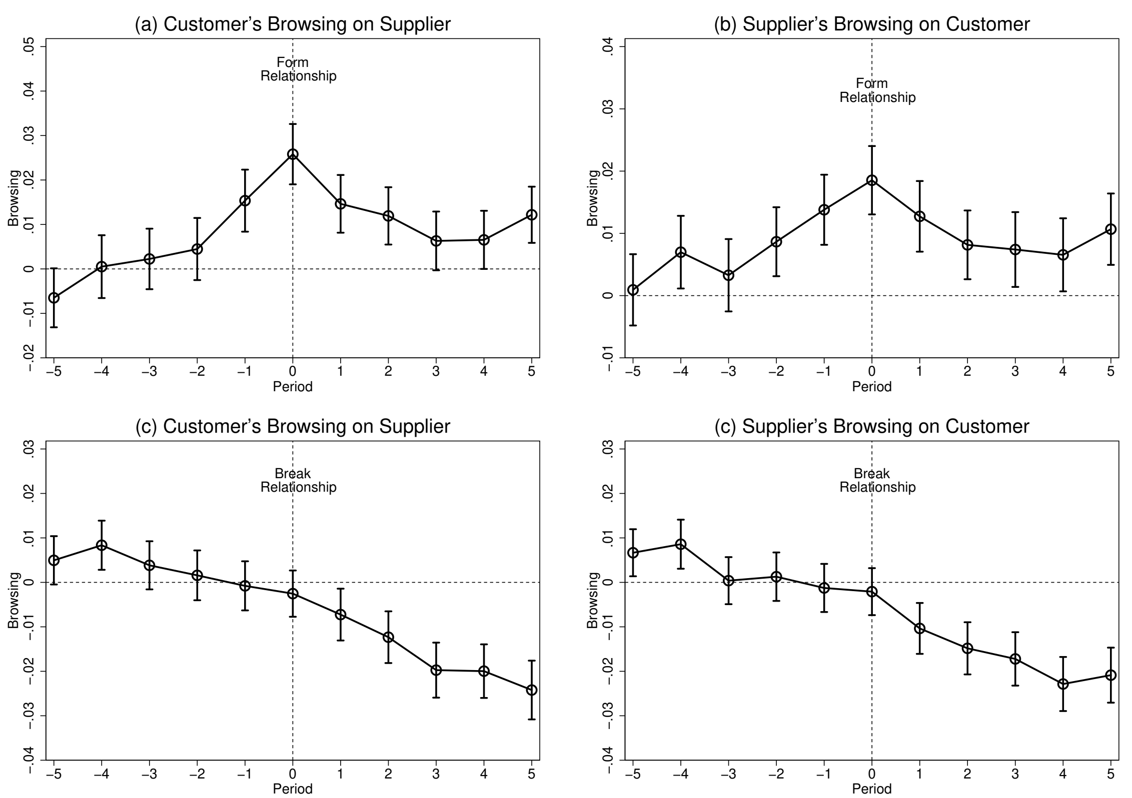
Note: The graph presents the number of suppliers/customers and supply chain relationships in the sample period.

Figure B.6: Browsing Intensity and Forecast Error



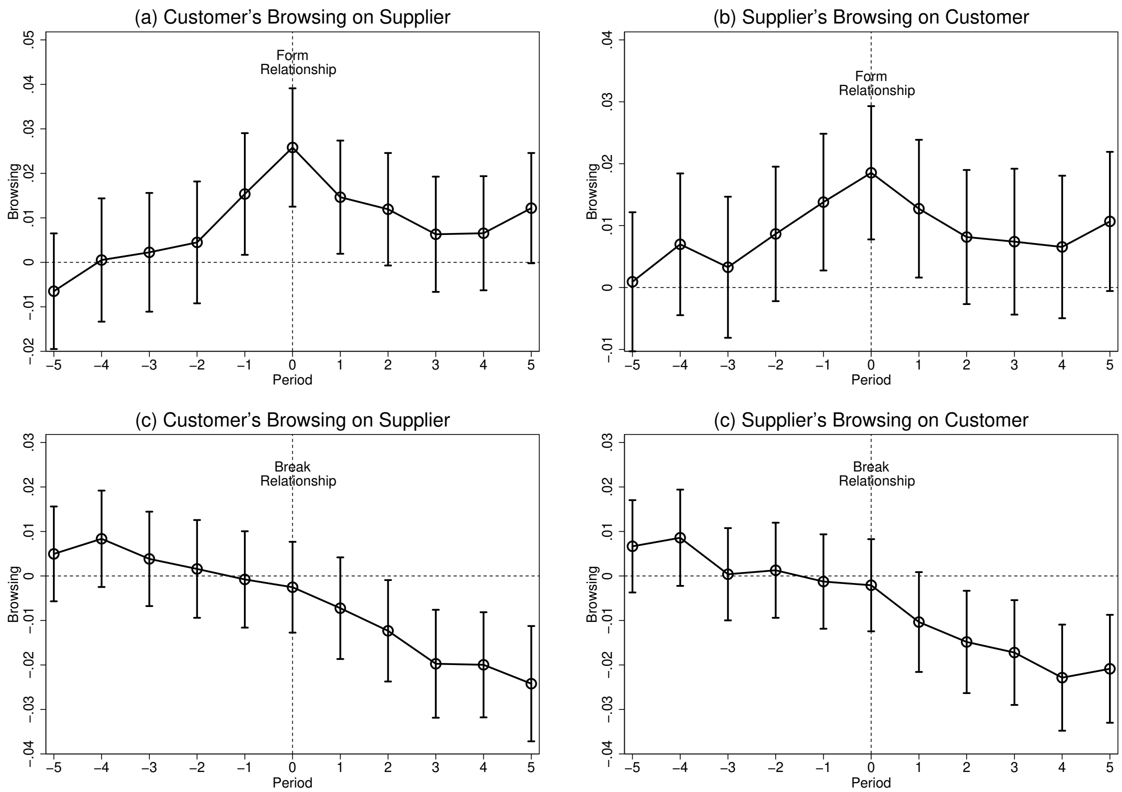
Note: The left panel is the binscatter plot between a firm's forecast error on its future earnings per share (EPS) and its total log browsing 90 days before the forecast. The right panel is the binscatter plot between a firm's forecast error on its future sales and its total log browsing 90 days before the forecast

Figure B.7: Browsing and Trading Relationship: Event Studies



Note: The figure depicts the log browsings of customers on suppliers (left panel) and suppliers on customers (right panel) when new trading relationships are formed at period 0. The frequency is quarterly and 90% confidence intervals are based on clustering at the browser level.

Figure B.8: Browsing and Trading Relationship: Event Studies (95 CI)

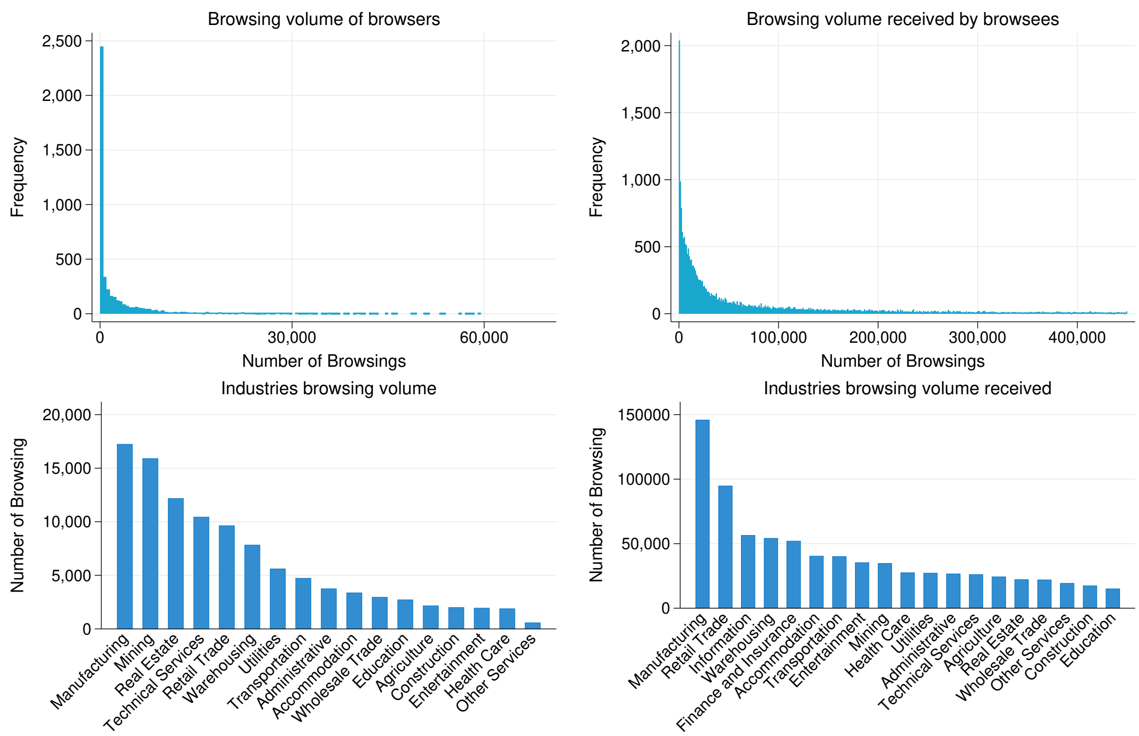


Note: The figure depicts the log browsings of customers on suppliers (left panel) and suppliers on customers (right panel) when new trading relationships are formed at period 0. The frequency is quarterly and 90% confidence intervals are based on clustering at the browser level.

C. Additional Facts on Browsing Activities

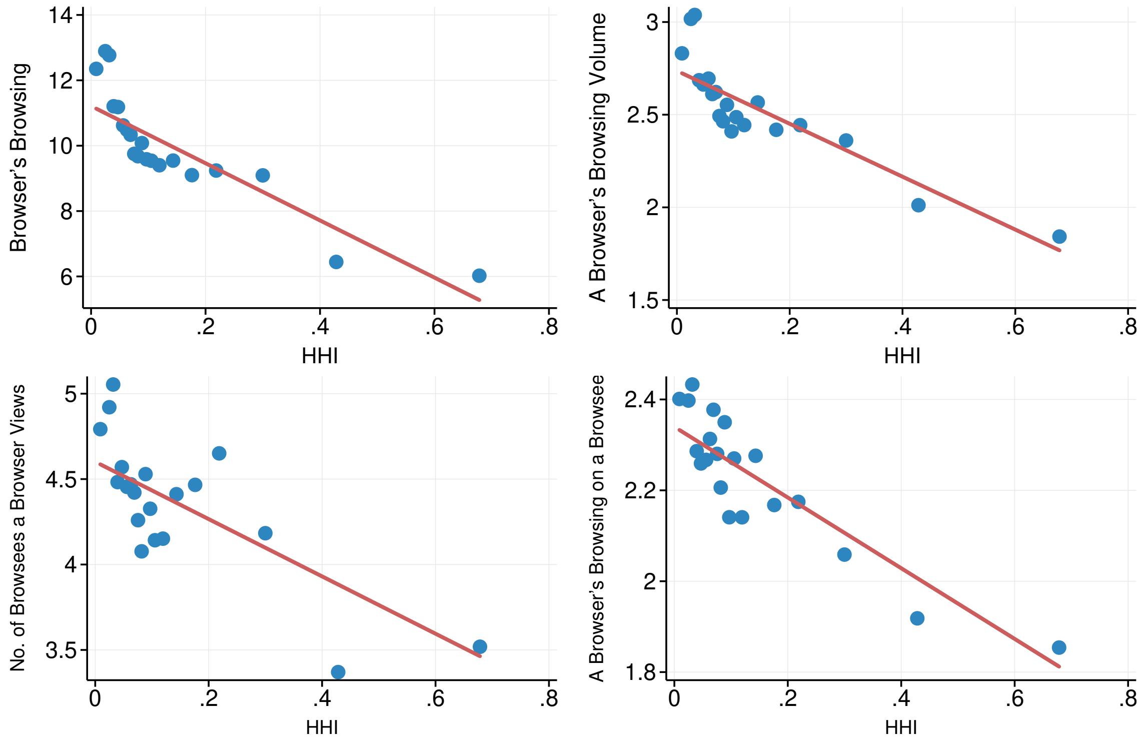
Distribution of browsing activities across firms and industries. Figure C.1 shows the distribution of browsing activity at the firm and industry level in 2016. To account for size bias, we adjust a firm's total browsing by its total employment (shown in Figure B.2) and by its sales (shown in Figure B.3), respectively. The browsings are distributed unequally across firms and industries. In particular, the distribution is fat-tailed and involves high kurtosis — a large fraction of firms have few browsings, while a small number of firms have significantly large browsings.

Figure C.1: Distribution of Browsing Activity



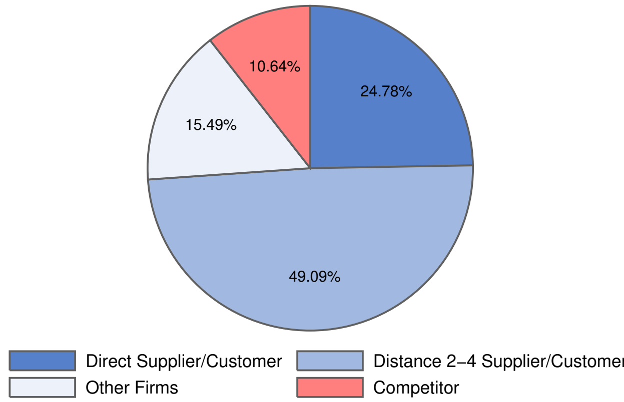
Browsing, employment, sales, and market concentration. Figure B.4 illustrates the relationship between a firm's total browsing with its employment (left panel) and sales (right panel). It is clear that larger firms, on average, browse other firms more intensively. Figure C.2 shows how browsers' browsings vary with the market concentration of the industry they are located in. We measure the market concentration using the Herfindahl-Hirschman Index (HHI). The y-axis in the top-left panel represents firms' total browsing volume in an industry. The y-axis in the top-right panel measures the average browsing volume of a browser in an industry. The y-axis in the bottom-left panel denotes the average number of browsees a browser views in an industry. Finally, the y-axis in the bottom-right panel is the average browsing on a browsee of a browser in an industry. To sum up, if a firm is located in an industry with a smaller HHI (or less concentrated), it, on average, has more total browsing. This is due to 1) it browses a larger number of firms, and 2) it also browses more intensively on each browsee.

Figure C.2: Browsing and Market Concentration



Attention allocated to competitors, suppliers and customers. Figure C.3 plots the allocation of a firm's browsing activity to three types of firms: the direct supplier or customer, its competitor, and firms that are neither direct suppliers/customers nor competitors. A typical firm allocates 24.78% of its browsing to its direct suppliers and customers, 49.09% to distance-2, distance-3 and distance-4 suppliers or customers, while it allocates 12.6% to its competitors. It is important to note that the attention allocated to direct and indirect suppliers and customers is possibly underestimated. The reasons are twofold. First, the Factset data only records a fraction of firms' suppliers and customers. Second, we classify a firm as a competitor if it is both a supplier/customer and a competitor. Therefore, our exercise provides a lower bound for quantifying firms' attention allocated to suppliers and customers.

Figure C.3: Allocation of Browsing Activity



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

D. Theory

In this paper, we adopt the convention that two vectors or matrices have the relation $\mathbf{A} \geq \mathbf{B}$ if $a_{ij} \geq b_{ij}, \forall i, j$, $\mathbf{A} > \mathbf{B}$ if $\mathbf{A} \geq \mathbf{B}$, and $\mathbf{A} \neq \mathbf{B}$. $\mathbf{A} >> \mathbf{B}$ if $a_{ij} > b_{ij}, \forall i, j$. Therefore $\mathbf{A} > \mathbf{0}$ implies that $a_{ij} \geq 0, \ \forall i, j$ and $\mathbf{A} \neq \mathbf{0}$. In addition, vectors in this paper are column vectors by default unless specified otherwise (for example, $\boldsymbol{\kappa}$ and $\boldsymbol{\varphi}$ are row vectors). Throughout our theoretical, quantitative, and optimal policy analysis, we impose the following parameter assumption on the economy's production network.

Assumption 1. *The economy's production network is represented by a nonnegative adjacency matrix $\mathbf{A} > \mathbf{0}$ such that for each sector $\sum_{j=1}^{N} a_{ij} < 1; \quad \forall i = 1, 2, ...N$. Each sector's production displays constant return to scale, therefore the labor share vector $\boldsymbol{\alpha} = \left(\{\alpha_i\}_{i=1}^{N}\right)'$ satisfies $\alpha_i = 1 - \sum_{j=1}^{N} a_{ij} > 0, \ \ \forall i = 1, 2, ...N$. We denote $'$ as the matrix transpose.*

In addition, in this paper we define the following matrix operator,

$$\mathrm{diag}(\cdot) : \mathbb{R}^{N\times N} \longmapsto \mathbb{R}^{N}; \qquad \mathrm{diag}(\cdot) : \mathbb{R}^{N} \longmapsto \mathbb{R}^{N\times N}$$

where operator diag(·) either extracts the diagonal vector from a given matrix as a column vector, or perform the inverse operation that transform a given vector into diagonal matrix with corresponding diagonal elements equal to the input vector. For example, consider a N-dimensional vector of interest ,

$$\boldsymbol{\mu}^{v} = \left(\mu_1, \mu_2, ...\mu_N\right)' \in \mathbb{R}^{N} = \mathrm{diag}\left(\boldsymbol{\mu}\right); \qquad \boldsymbol{\mu} = \mathrm{diag}\left(\boldsymbol{\mu}^{v}\right)$$

Therefore, the nature of this matrix operation hinges on the input object, which is consistent with the Matlab convention.

Next, we adopt the following conventions for matrix differentiation:

1. The derivative between a scalar and a vector follows an arrangement of the same form as this vector. For example,
$$\frac{\partial \mu_i}{\partial \boldsymbol{\kappa}} = \begin{bmatrix} \frac{\partial \mu_i}{\partial \kappa_1} & \frac{\partial \mu_i}{\partial \kappa_2} & \cdots & \frac{\partial \mu_i}{\partial \kappa_N} \end{bmatrix}; \qquad \frac{\partial \boldsymbol{\mu}^{v}}{\partial \kappa_i} = \begin{bmatrix} \frac{\partial \mu_1}{\partial \kappa_i} & \frac{\partial \mu_2}{\partial \kappa_i} & \cdots & \frac{\partial \mu_N}{\partial \kappa_i} \end{bmatrix}'$$
2. The derivative between a scalar and a matrix follows an arrangement of the same form as this matrix. For example,
$$\frac{\partial \mu_i}{\partial \mathbf{A}} = \begin{bmatrix} \frac{\partial \mu_i}{\partial a_{11}} & \frac{\partial \mu_i}{\partial a_{12}} & \cdots & \frac{\partial \mu_i}{\partial a_{1N}} \\ \frac{\partial \mu_i}{\partial a_{21}} & \frac{\partial \mu_i}{\partial a_{22}} & \cdots & \frac{\partial \mu_i}{\partial a_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mu_i}{\partial a_{N1}} & \frac{\partial \mu_i}{\partial a_{N2}} & \cdots & \frac{\partial \mu_i}{\partial a_{NN}} \end{bmatrix}; \qquad \frac{\partial \mathbf{A}}{\partial \mu_i} = \begin{bmatrix} \frac{\partial a_{11}}{\partial \mu_i} & \frac{\partial a_{12}}{\partial \mu_i} & \cdots & \frac{\partial a_{1N}}{\partial \mu_i} \\ \frac{\partial a_{21}}{\partial \mu_i} & \frac{\partial a_{22}}{\partial \mu_i} & \cdots & \frac{\partial a_{2N}}{\partial \mu_i} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial a_{N1}}{\partial \mu_i} & \frac{\partial a_{N2}}{\partial \mu_i} & \cdots & \frac{\partial a_{NN}}{\partial \mu_i} \end{bmatrix}$$
3. The derivative between two vectors: the numerator vector determines the arrangement of its derivative

matrix. For example,

$$\frac{\partial \boldsymbol{\mu}^v}{\partial \boldsymbol{\kappa}} = \frac{\partial \boldsymbol{\mu}^v}{\partial \boldsymbol{\kappa}'} = \begin{bmatrix} \frac{\partial \boldsymbol{\mu}^v}{\partial \kappa_1} & \frac{\partial \boldsymbol{\mu}^v}{\partial \kappa_2} & \cdots & \frac{\partial \boldsymbol{\mu}^v}{\partial \kappa_N} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mu_1}{\partial \kappa_1} & \frac{\partial \mu_1}{\partial \kappa_2} & \cdots & \frac{\partial \mu_1}{\partial \kappa_N} \\ \frac{\partial \mu_2}{\partial \kappa_1} & \frac{\partial \mu_2}{\partial \kappa_2} & \cdots & \frac{\partial \mu_2}{\partial \kappa_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mu_N}{\partial \kappa_1} & \frac{\partial \mu_N}{\partial \kappa_2} & \cdots & \frac{\partial \mu_N}{\partial \kappa_N} \end{bmatrix}; \qquad \frac{\partial (\boldsymbol{\mu}^v)'}{\partial \boldsymbol{\kappa}} = \frac{\partial (\boldsymbol{\mu}^v)'}{\partial \boldsymbol{\kappa}'} = \begin{bmatrix} \frac{\partial \mu_1}{\partial \kappa_1} & \frac{\partial \mu_2}{\partial \kappa_1} & \cdots & \frac{\partial \mu_N}{\partial \kappa_1} \\ \frac{\partial \mu_1}{\partial \kappa_2} & \frac{\partial \mu_2}{\partial \kappa_2} & \cdots & \frac{\partial \mu_N}{\partial \kappa_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mu_1}{\partial \kappa_N} & \frac{\partial \mu_2}{\partial \kappa_N} & \cdots & \frac{\partial \mu_N}{\partial \kappa_N} \end{bmatrix}$$

D.1 Proof of Proposition 3.1

Proof. In model equilibrium, households maximize their expected utility subject to their budget constraints. The optimal consumption demand and the intratemporal Euler condition satisfy

$$P_{i,t} C_{i,t} = \beta_i P_t C_t; \qquad \frac{W_t}{P_t} C_t^{-\gamma} = L_t^{\frac{1}{\eta}}$$

With perfect information, firms' profit optimization problem reads

$$\max_{\left\{P_{i\iota t}, L_{i\iota t}, \{X_{i\iota jt}\}_{j=1}^N\right\}} \Pi_{i\iota t} = (1+\tau_i) P_{i\iota t} Y_{i\iota t} - W_t L_{i\iota t} - \sum_{j=1}^N P_{jt} X_{i\iota jt} - T_{i\iota t}$$

where $\tau_i = \frac{1}{(\theta_i - 1)}$ is the subsidy to remove monopolistic distortion, $T_{i\iota t}$ is a lump-sum tax that finances the subsidy. In equilibrium, $T_{i\iota t} = \tau_i P_{i\iota t} Y_{i\iota t}$. A canonical cost minimization problem implies that the optimal input and labor demand satisfy

$$X_{i,\iota,j,t} = a_{ij} \mathrm{MC}_{i,t} \frac{Y_{i,\iota,t}}{P_{j,t}}, \quad \text{and} \quad L_{i,t} = \alpha_i \mathrm{MC}_{i,t} \frac{Y_{i,\iota,t}}{W_t}, \tag{D.1}$$

where the marginal cost of firms in sector i is $\mathrm{MC}_{it} = Z_{it}^{-1} W_t^{\alpha_i} \prod_{j=1}^N P_{jt}^{a_{ij}}$. The market-clearing condition is given by

$$Y_{i,t} = C_{i,t} + \sum_{j=1}^N \int_0^1 X_{j,\iota,i,t} d\iota.$$

Next, we characterize the full-information equilibrium.

Frictionless benchmark. In absence of information acquisition cost, the sectoral price is identical to the nominal marginal cost,

$$P_{i,\iota,t} = P_{i,t} = \mathrm{MC}_{i,t} = \frac{1}{Z_{i,t}} W_t^{\alpha_i} \prod_{j=1}^N P_{j,t}^{a_{ij}}, \tag{D.2}$$

As a result, the after-tax (net) profit for each firm equals to zero and the households' budget balance is given by

$P_tC_t = W_tL_t$. And the sectoral inputs follow

$$X_{j,i,t} = X_{j,\iota,i,t} = a_{ji}\frac{P_{j,t}Y_{j,\iota,t}}{P_{i,t}}. \tag{D.3}$$

Then the goods market clearing condition for sector i follows

$$P_{i,t}Y_{i,t} = \beta_i P_t C_t + \sum_{j=1}^{N} a_{ji}P_{j,t}Y_{j,t}. \tag{D.4}$$

Let λ_{it} denote the sales-to-GDP ratio (the Domar weight), thus

$$\lambda_{i,t} = \beta_i + \sum_{j=1}^{N} a_{ji}\lambda_{j,t}.$$

Consequently, with perfect information, the Domar weight is constant. In matrix form,

$$\boldsymbol{\lambda} = \boldsymbol{\beta} + \mathbf{A}'\boldsymbol{\lambda}.$$

Then the Domar weight is given by

$$\boldsymbol{\lambda}' = \boldsymbol{\beta}'(\mathbf{I} - \mathbf{A})^{-1}. \tag{D.5}$$

Next, we provide the steady-state solution of the frictionless benchmark.

The determinant steady-state. The subscript t of a time-varying variable is removed to denote the steady-state value. We normalize the steady-state money supply M to be 1, which implies that the GDP is equal to 1 as $PC = M = 1$. In the perfect-information economy, $P_{i,\iota} = P_i = \mathrm{MC}_i$, so the steady-state value of the prices and wage rate in (D.2) satisfy

$$P_i = \mathrm{MC}_i = \frac{1}{Z_i}W^{\alpha_i}\prod_{j=1}^{N}P_j^{a_{ij}} = W^{\alpha_i}\prod_{j=1}^{N}P_j^{a_{ij}}, \tag{D.6}$$

where the steady-state of sectoral productivity shock $Z_i = 1$, $i = 1, 2, ...N$. Meanwhile, each firm's after-tax profit is zero, $\Pi_{i,\iota} = 0$. The steady-state household budget constraint is given by $PC = WL = M$, and the optimal labor supply condition follows

$$\frac{W}{P}C^{-\gamma} = L^{\frac{1}{\eta}}.$$

Combing these two equations.

$$W = M^{\frac{1+\eta\gamma}{1+\eta}} P^{\frac{\eta(1-\gamma)}{1+\eta}} = M^{\frac{1+\eta\gamma}{1+\eta}} \left(\prod_{j=1}^{N} P_j \right)^{\frac{\beta_j \eta(1-\gamma)}{1+\eta}}, \tag{D.7}$$

where we use $P = \prod_{i=1}^{N} P_i^{\beta_i}$ and rewrite the sector's subscript to j. Given equation (D.6) and (D.7), the steady-state of equilibrium price system is given by

$$P_i = M^{\frac{\alpha_i(1+\eta\gamma)}{1+\eta}} \left(\prod_{j=1}^{N} P_j \right)^{\frac{\alpha_i \beta_j \eta(1-\gamma)}{1+\eta} + a_{ij}},$$

which implies that the steady-state of sectors' prices are determined by the exogenous money supply $M = 1$. Explicitly, $P_i = 1$ for $i = 1, 2, ...N$. Thus, the final goods price P and the final consumption goods C can be obtained in turn:

$$P = \prod_{i=1}^{N} P_i^{\beta_i} = 1, \quad \text{and} \quad C = \frac{M}{P} = 1,$$

Furthermore, the steady-state wage rate and the aggregated labor is given by

$$W = 1, \quad \text{and} \quad L = \frac{M}{W} = 1.$$

Next, since $P_i C_i = \beta_i P C$, the sectoral consumption follows

$$C_i = \frac{\beta_i P C}{P_i} = \beta_i.$$

Using (D.1) under perfect information, firms' inputs in sector i inputs are given by

$$L_i = \alpha_i \frac{P_i Y_i}{W} = \alpha_i \lambda_i, \quad \text{and} \quad X_{i,j} = X_{i,\iota,j} = a_{ij} \frac{P_i Y_i}{P_j} = a_{ij} \lambda_i$$

Finally, sector i's steady-state output is given by

$$Y_i = \frac{\lambda_i M}{P_i} = \lambda_i$$

Next, we log-linearize the equilibrium system under perfect information.

Linearized system under perfect information. The log-linearized goods market clearing condition (D.4) is given by

$$P_i Y_i \left(p_{it} + y_{it} \right) = \beta_i M m_t + \sum_{j=1}^{N} a_{ji} P_j Y_j \left(p_{jt} + y_{jt} \right)$$

where we use $P_t C_t = M_t$ and $P_{it} = MC_{it}$. , In matrix form:

$$\boldsymbol{p}_t + \boldsymbol{y}_t = \text{diag}\,(\boldsymbol{\lambda})^{-1}\,(\mathbf{I} - \mathbf{A}')^{-1}\boldsymbol{\beta} m_t \tag{D.8}$$

Next, the sectoral price in (D.2) can be log-linearized as

$$p_{it} = \text{mc}_{it} = -z_{it} + \alpha_i w_t + \sum_j a_{ij} p_{jt}.$$

To facilitate comparison with the *Rational-Inattention equilibrium*, we denote $\boldsymbol{p}_t^*$ and $\boldsymbol{c}_t^*$ as the price-response vector and consumption vector in the *Full-Information Equilibrium* , respectively. The matrix form of the above equation can be expressed as

$$\boldsymbol{p}_t^* = -\boldsymbol{z}_t + \boldsymbol{\alpha} w_t + \mathbf{A}\boldsymbol{p}_t^*,$$

which leads to the price-response vector under perfect information,

$$\boldsymbol{p}_t^* = (\mathbf{I} - \mathbf{A})^{-1}(-\boldsymbol{z}_t + \boldsymbol{\alpha} w_t). \tag{D.9}$$

We define p_t^f as the log-deviation from the final goods price P_t , then $p_t^f = \sum_{i=1}^N \beta_i p_{it} = \boldsymbol{\beta}'\boldsymbol{p}_t^*$. Multiplying both sides of the equation (D.9) by $\boldsymbol{\beta}'$ yields

$$p_t^f = -\boldsymbol{\beta}'(\mathbf{I} - \mathbf{A})^{-1}\boldsymbol{z}_t + \boldsymbol{\beta}'(\mathbf{I} - \mathbf{A})^{-1}\boldsymbol{\alpha} w_t = -\boldsymbol{\lambda}'\boldsymbol{z}_t + w_t, \tag{D.10}$$

where we use the matrix property of $\boldsymbol{\beta}'(\mathbf{I} - \mathbf{A})^{-1}\boldsymbol{\alpha} = 1$ and $\boldsymbol{\beta}'(\mathbf{I} - \mathbf{A})^{-1} = \boldsymbol{\lambda}'$. We then log-linearize the households' budget constraint under perfect information and the Euler condition, yielding

$$p_t^f + c_t^* = w_t + \ell_t = m_t, \quad \text{and} \quad w_t^f - p_t^f - \gamma c_t^* = \frac{1}{\eta}\ell_t.$$

Eliminating ℓ_t by combining the above equations provides the linearized output as

$$c_t^* = \frac{1+\eta}{1+\gamma\eta}\left(w_t - p_t^f\right) = \frac{1+\eta}{1+\gamma\eta}\boldsymbol{\lambda}'\boldsymbol{z}_t,$$

where the second identity follows from the equation (D.10).

□

D.2 Proof of Lemma 3.1

Proof. We first derive a linear-quadratic-Gaussian (LQG) approximation to each firm's expected profit function. We focus on a firm ι from sector i, at the first stage, its expected profit function is given by

$$\max_{P_{i\iota t}} \mathbb{E}_{i\iota t}\left[\frac{U'(C_t)}{P_t}\left\{\left((1+\tau_i)P_{i\iota t}^{(1-\theta_i)} - P_{i\iota t}^{-\theta_i}MC_{it}\right)P_{it}^{\theta_i}Y_{it}\right\}\right],$$

where we substitute firm (i,ι)'s demand function under monoplistic competition into the objective. Since the representative household's stochastic discount factor $\frac{U'(C_t)}{P_t}$ and does not affect firms' optimal pricing decisions in our static model, they do not affect firms' optimal information choices. Therefore, we simplify the objective function as[29]

$$\max_{P_{i\iota t}} \mathbb{E}_{i\iota t}\Bigg[\underbrace{(1+\tau_i)P_{i\iota t}^{(1-\theta_i)}P_{it}^{\theta_i}Y_{it} - P_{i\iota t}^{-\theta_i}P_{it}^{\theta_i}Y_{it}MC_{it}}_{\equiv \Pi_{i\iota t}}\Bigg],$$

which parallels (3.1). Its marginal cost is given by

$$MC_{it} = \frac{W_t^{\alpha_i}\prod_{j=1}^{N}P_{jt}^{a_{ij}}}{Z_{it}}.$$

Note that the marginal cost is identical across firms within the same sector. In preceding analysis, we present the solution of perfect-information, deterministic steady state in this model, so we express the objective function in log-deviations from the steady-state,

$$\Pi_{i\iota t} = \frac{\theta_i}{(\theta_i-1)}\lambda_i e^{\left[(1-\theta_i)p_{i\iota t}+\theta_i p_{it}+y_{it}\right]} - \lambda_i e^{\left[-\theta_i p_{i\iota t}+\theta_i p_{it}+y_{it}+\alpha_i w_t+\sum_{j=1}^{N}a_{ij}p_{jt}-z_{it}\right]} \tag{D.11}$$

To ease the exposition, we define a vector of state variables for firm (i,ι),

$$\boldsymbol{\Upsilon}_{it} = \begin{bmatrix} p_{1t} & p_{2t} & \ldots, p_{it} & \ldots & p_{Nt} & w_t & z_{it} & y_{it}\end{bmatrix}'.$$

Next, we perform a second-oder Taylor series expansion of $\Pi_{i\iota t}$ in terms of the control variable $p_{i\iota t}$ and the state vector $\boldsymbol{\Upsilon}_{it}$,

$$\Pi_{i\iota t}\left(p_{i\iota t},\boldsymbol{\Upsilon}_{it}\right) = \Pi_{i\iota t}(0,\mathbf{0}) + \mathbf{D}_{\Pi_{i\iota}(0,\mathbf{0})}\mathcal{X}_{i\iota t} + \frac{1}{2}\mathcal{X}'_{i\iota t}\mathbf{H}_{\Pi_{i\iota}(0,\mathbf{0})}\mathcal{X}_{i\iota t} + O(\varepsilon^3). \tag{D.12}$$

where $\mathcal{X}_{i\iota t} \equiv \begin{bmatrix}\boldsymbol{\Upsilon}_{it}\\ p_{i\iota t}\end{bmatrix}$, $\mathbf{D}_{\Pi_{i\iota}(0,\mathbf{0})}$ and $\mathbf{H}_{\Pi_{i\iota}(0,\mathbf{0})}$ are Jacobian and Hessian matrix of derivatives, respectively. Using standard LQG control techniques, it is straightforward to show that only the quadratic term matters for the

[29] The term $P_{it}^{\theta_i}Y_{it}$ can also be removed from the objective, because the LQG approximation will be unique up to the steady-state constant $\frac{U'(C)}{P}P_i^{\theta_i}Y_i$, but such difference can be normalized by choosing approriate information cost χ_i.

optimal information choice under rational inattention.[30] In the approximation, the Hessian matrix is given by $\frac{1}{2}\mathbf{H}_{\Pi_{it}(0,\mathbf{0})} \equiv \begin{bmatrix} \mathbf{H}^i_{xx} & (\mathbf{H}^i_{ux})' \\ \mathbf{H}^i_{ux} & H^i_{uu} \end{bmatrix}$, where

$$H^i_{uu} = -\frac{1}{2}\lambda_i\theta_i; \qquad \mathbf{H}^i_{ux} = \frac{1}{2}\lambda_i\theta_i \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{iN} & \alpha_i & -1 & 0 \end{bmatrix}.$$

We do not spell the expression for block matrix $\mathbf{H}^i_{xx}$ because it is irrelevant for the optimal pricing decision and information choice. Then, we express the second-order approximation of the profit function under rational inattention in quadratic form

$$\max_{p_{iit}} -\mathbb{E}_{iit}\left[\boldsymbol{\Upsilon}'_{it}\mathbf{W}_i\boldsymbol{\Upsilon}_{it} + p'_{iit}R_i p_{iit} + 2\boldsymbol{\Upsilon}'_{it}\mathbf{S}_i p_{iit}\right] \tag{D.13}$$

with coefficient matrices defined by

$$\mathbf{W}_i = -\mathbf{H}^i_{xx}; \quad R_i = -H^i_{uu}; \quad \mathbf{S}_i = -(\mathbf{H}^i_{ux})'.$$

such that $\begin{bmatrix} \mathbf{W}_i & \mathbf{S}_i \\ \mathbf{S}'_i & R_i \end{bmatrix} \succeq 0$ and $R_i \succ 0$. Using acquisition-filter-control separation principle and the certainty-equivalence principle of the LQG control Tanaka et al. (2017), the optimal pricing decision is given by[31]

$$p_{iit} = -\mathbf{F}_i\mathbb{E}_{iit}\left[\Upsilon_{it} \mid \boldsymbol{x}_{iit}\right] = \mathbb{E}_{iit}\left[\mathrm{mc}_{it} \mid \boldsymbol{x}_{iit}\right]; \quad \mathbf{F}_i = R_i^{-1}\mathbf{S}'_i = -\begin{bmatrix} \mathbf{e}_i\mathbf{A} & \alpha_i & -1 & 0 \end{bmatrix}. \tag{D.14}$$

Agggating (D.14) over the the entire sector, the linear functions of prices and marginal costs coincides with log-linear decision rule (3.3). The equivalence between log-linear approximation and linear-quadratic approximation remains valid for ANY information structure firm chooses. Since y_{it} does not appear in firm's ideal price, the marginal cost, it has no incentive to pay attention to this state variable. Therefore, from now on we drop y_{it} and refine the state variable as $\boldsymbol{\Upsilon}_{it} = \begin{bmatrix} p_{1t} & p_{2t} & \dots, p_{it} & \dots & p_{Nt} & w_t & z_{it} \end{bmatrix}'$. The Hessian block matrix is given by $\mathbf{H}^i_{ux} = \frac{1}{2}\lambda_i\theta_i \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{iN} & \alpha_i & -1 \end{bmatrix}$. The optimal decision rule is simplified to $\mathbf{F}_i = -\begin{bmatrix} \mathbf{e}_i\mathbf{A} & \alpha_i & -1 \end{bmatrix}$.[32] Next, we employ Lemma 1 in Miao et al. (2022) to cast the profit maximization problem of (D.13) into a tracking problem of information choices,

$$\max_{\boldsymbol{x}_{iit}} -\mathbb{E}\left[\left(\boldsymbol{\Upsilon}_{it} - \mathbb{E}\left[\boldsymbol{\Upsilon}_{it} \mid \boldsymbol{x}_{iit}\right]\right)' \boldsymbol{\Omega}_i \left(\boldsymbol{\Upsilon}_{it} - \mathbb{E}\left[\boldsymbol{\Upsilon}_{it} \mid \boldsymbol{x}_{iit}\right]\right)\right] \tag{D.15}$$

where the positive-semidefinite weighting matrix is given by $\boldsymbol{\Omega}_i = \mathbf{F}'_i R_i \mathbf{F}_i \succeq 0$. Using this observation and (D.14),

[30] To see this point, we can transform (D.12) into a pure quadratic form (D.13) by introducing an additional constant state 1, leading to a augmented tracking problem akin to (D.15). Since all firms are able to forecast the deterministic state 1 perfectly without incuring any cost, its associated mean-square errors (variance and covariances) are identically 0. Therefore, we can safely disregard the constant and first-order term in the RI problem.

[31] We denote $\mathbf{e}_i$ as the ith standard basis row vector in $\mathbb{R}^N$.

[32] It is straightforward to verify that such refinement doe not change any of our results.

we obtain the tracking formulation in terms of optimal pricing decision,

$$\max_{\boldsymbol{x}_{i\iota t}} -\frac{1}{2}\lambda_i\theta_i\mathbb{E}\left[\left(p_{i\iota t}-\mathrm{mc}_{it}\right)^2\right]$$

as desired. The existence of a pure tracking problem in optimal information choice relies on the fact the state variables Υ_{it} are "exogenous" in the eyes of firm (i,ι). Since information costs χ are sector-specific, it is clear that firms' optimal information choice problems are homogeneous within each sector but heterogeneous across sectors in the network. In the tracking formulation, firm's information acquisition motive is dictated by the optimal pricing decision, which in turn hinge on network structure via the marginal cost.

Finally, we note that the pricing decision and state variables $\boldsymbol{\Upsilon}_{it}$ are both functions of sectoral productivity shocks $\boldsymbol{z}_t$. Under LQG setting with Shannon-entropy cost function $C_i(\boldsymbol{x}_{i,\iota,t})$, it is well known that the optimal information structure is Gaussian, represented by an signal structure in the form

$$\boldsymbol{x}_{i,\iota,t}=\mathbf{H}_i\boldsymbol{z}_t+\mathbf{u}_{i,l,t},\quad \mathbf{u}_{i,\iota,t}\sim\mathbb{N}\left(\mathbf{0},\mathbf{V}_i\right)$$

where $\mathbf{H}_i$ and $\mathbf{V}_i\succ 0$ are matrices of unknown dimensions to be determined by the RI problem. The information cost function is parameterized by the classical Gaussian (conditional) mutual information, $\mathbb{I}\left((\boldsymbol{z}_t;\boldsymbol{x}_{i\iota t}|\boldsymbol{\Sigma}_z\right)=\frac{1}{2}\left(\log\det\boldsymbol{\Sigma}_z-\log\det\boldsymbol{\Sigma}_{z|x_i}\right)$. The proof is now complete. □

D.3 Proof of Proposition 3.2

Proof. We first transform the optimal information choice problem into a semidefinite progamming problem. Let $\boldsymbol{p}_t=\boldsymbol{\phi}\boldsymbol{z}_t$. $\boldsymbol{\phi}$ is the pricing function ($N\times N$ influence matrix) under rational inattention, which is determined in general equilibrium.

Lemma D.1. *The RI problem for firms in sector i is represented by the following matrix optimization problem,*

$$\begin{aligned}\min_{\boldsymbol{\Sigma}_{z|x_i}}\ &\mathrm{tr}\left(\boldsymbol{\Omega}_{iz}\boldsymbol{\Sigma}_{z|x_i}\right)+\chi_i\left[\frac{1}{2}\log\det\boldsymbol{\Sigma}_z-\frac{1}{2}\log\det\boldsymbol{\Sigma}_{z|x_i}\right]\\ s.t.\quad &\mathbf{0}\preceq\boldsymbol{\Sigma}_{z|x_i}\preceq\boldsymbol{\Sigma}_z,\end{aligned}$$

where $\boldsymbol{\Sigma}_{z|x_i}$ denotes the posterior covariance matrices of the fundamental shocks $\boldsymbol{z}_t$. The modified weighting matrix admits representation as $\boldsymbol{\Omega}_{iz}=\mathbf{G}_i'\mathbf{G}_i$, where $\mathbf{G}_i=\sqrt{R_i}\mathbf{e}_i\left(\mathbf{A}\boldsymbol{\phi}-\mathbf{I}+\boldsymbol{\alpha\kappa}\right)$.

Proof. We write (3.3) in matrix form as

$$\mathrm{mc}_{it}=p^{\Delta}_{i\iota t}=\mathbf{e}_i\left(-\mathbf{I}+\boldsymbol{\alpha\kappa}+\mathbf{A}\boldsymbol{\phi}\right)\boldsymbol{z}_t, \tag{D.16}$$

where $p^{\Delta}_{i\iota t}$ denotes profit-maximizing ideal price. Substite (D.14) and (D.16) into the objective function in Lemma

3.1, we express the tracking problems in terms of $\boldsymbol{z}_t$,[33]

$$\min_{\boldsymbol{x}_{iit}} \mathrm{E}\left[(\boldsymbol{z}_t - \mathrm{E}[\boldsymbol{z}_t|\boldsymbol{x}_{iit}])' \boldsymbol{\Omega}_{iz} (\boldsymbol{z}_t - \mathrm{E}[\boldsymbol{z}_t|\boldsymbol{x}_{iit}])\right] + \chi_i \mathbb{I}(\boldsymbol{z}_t; \boldsymbol{x}_{iit}|\Sigma_{\boldsymbol{z}}), \tag{D.17}$$

where the modified weight matrix is given by $\boldsymbol{\Omega}_{iz} = \mathbf{G}_i'\mathbf{G}_i \geq 0$, where $\mathbf{G}_i = \sqrt{R_i}\mathbf{e}_i (\mathbf{A}\boldsymbol{\phi} - \mathbf{I} + \alpha\boldsymbol{\kappa})$. By basic matrix algebra, the first term in problem (D.17) can be expressed as

$$\begin{aligned}
\mathrm{E}\left[(\boldsymbol{z}_t - \mathrm{E}[\boldsymbol{z}_t|\boldsymbol{x}_{iit}])' \boldsymbol{\Omega}_{iz} (\boldsymbol{z}_t - \mathrm{E}[\boldsymbol{z}_t|\boldsymbol{x}_{iit}])\right] &= \mathrm{E}\left[\mathrm{tr}\left((\boldsymbol{z}_t - \mathrm{E}[\boldsymbol{z}_t|\boldsymbol{x}_{iit}])' \boldsymbol{\Omega}_{iz} (\boldsymbol{z}_t - \mathrm{E}[\boldsymbol{z}_t|\boldsymbol{x}_{iit}])\right)\right] \\
&= \mathrm{E}\left[\mathrm{tr}\left(\boldsymbol{\Omega}_{iz} (\boldsymbol{z}_t - \mathrm{E}[\boldsymbol{z}_t|\boldsymbol{x}_{iit}])(\boldsymbol{z}_t - \mathrm{E}[\boldsymbol{z}_t|\boldsymbol{x}_{iit}])'\right)\right] \\
&= \mathrm{tr}\left(\boldsymbol{\Omega}_{iz}\boldsymbol{\Sigma}_{z|x_i}\right),
\end{aligned}$$

where the second equality follows from the invariance property that the trace operator under circular shifts. Finally, given any signal structure firms choose, they are fully rational in predicting the unobserved states. Therefore, the Bayesian updating formula holds such that

$$\mathrm{SNR}_i \equiv \mathbf{H}_i'\mathbf{V}_i^{-1}\mathbf{H}_i = \boldsymbol{\Sigma}_{z|x_i}^{-1} - \boldsymbol{\Sigma}_z^{-1} \geq 0. \tag{D.18}$$

where SNR_i is defined as matrix-valued signal-to noise ratio. Since $\boldsymbol{\Sigma}_{z|x_i} \geq 0$ is by construction positive-semidefinite, elementary theory of positive-definite matrices implies that

$$0 \preceq \boldsymbol{\Sigma}_{z|x_i} \preceq \boldsymbol{\Sigma}_z$$

The underlying "no-forgetting constraint" reflects the full-rationality of firms: they cannot achieve better forecasts of some states by eliminating existing memories of other states. This completes the proof. □

Next, we derive a closed-form solution to the matrix optimization problem defined in Lemma D.1. The solution then produces the optimal signal structure as a by-product. In particular, we employ techniques developed by Miao et al. (2022), which also apply to more general models of dynamic, multivariate rational inattention.

To begin with, we note that by (D.16), $\mathrm{mc}_{it} = p_{it}^{\Delta} = R_i^{-\frac{1}{2}}\mathbf{G}_i\boldsymbol{z}_t$. We define the variance of the marginal cost (profit-maximizing price, or ideal price) as $\mathbb{V}(\mathrm{mc}_{it})$. As we emphasized in the main text, this marginal-cost volatility serves as the key determinant of firms' information acquisition strategy. The following relationship between $\mathbb{V}(\mathrm{mc}_{it})$ and $\mathbf{G}_i$ holds:

$$\mathbb{V}(\mathrm{mc}_{it}) = R_i^{-1} \parallel \boldsymbol{\Sigma}_z^{\frac{1}{2}}\mathbf{G}_i' \parallel^2, \tag{D.19}$$

where $\parallel \cdot \parallel$ denotes the Euclidean norm. Since $\boldsymbol{\Omega}_{iz}$ is symmetric and $\boldsymbol{\Sigma}_z > 0$, we construct a symmetric, positive-

[33] Alternatively, we can use (D.15) to derive (D.17) by defining the mapping betweem endogenous states and shocks: $\Upsilon_{it} = \boldsymbol{\Psi}_i\boldsymbol{z}_t$. $\boldsymbol{\Psi}_i = \begin{bmatrix}\boldsymbol{\phi} \\ \boldsymbol{\kappa} \\ \mathbf{e}_i\end{bmatrix}$. Two derivations are equivalent.

semidefinite matrix $\mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{\Omega}_{iz}\mathbf{\Sigma}_z^{\frac{1}{2}}$, where $\mathbf{\Sigma}_z^{\frac{1}{2}} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma_N \end{bmatrix} \succ 0$ denotes the positive square root matrix of the prior covariance matrix. Next, perform an unitary eigendecomposition,

$$\mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{\Omega}_{iz}\mathbf{\Sigma}_z^{\frac{1}{2}} = \mathbf{U}_i\mathbf{\Omega}_d^i\mathbf{U}_i',$$

where $\mathbf{U}_i\mathbf{U}_i' = I$ is unitary and $\mathbf{\Omega}_d^i = \operatorname{diag}\left(\{d_k^i\}_{k=1}^N\right)$ is a diagonal matrix of nonnegative eigenvalues with descending order. Note that $\mathbf{G}_i = \sqrt{R_i}\mathbf{e}_i\left(\mathbf{A}\boldsymbol{\phi} - \mathbf{I} + \boldsymbol{\alpha\kappa}\right)$ is a $1 \times N$ vector, and the rank of the outer product $\mathbf{\Omega}_{iz} = \mathbf{G}_i'\mathbf{G}_i$ is equal to one. Therefore, $\operatorname{rank}\left(\mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{\Omega}_{iz}\mathbf{\Sigma}_z^{\frac{1}{2}}\right) = 1$ by construction. Consequently, there is only one strictly positive eigenvalue with an associated eigenvector denoted as

$$\begin{aligned} d_1^i &= \| \mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{G}_i' \|^2 = \frac{1}{2}\theta_i\lambda_i\mathbb{V}\left(\mathrm{mc}_{it}\right), \\ \zeta_1^i &= \frac{\mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{G}_i'}{\| \mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{G}_i' \|}. \end{aligned} \tag{D.20}$$

The rest of eigenvalues are equal to 0 by definition. The next lemma presents a generalized reverse water-filling solution of optimal posterior matrix $\mathbf{\Sigma}_{z|x_i}$,[34]

Lemma D.2. *The optimal solution to the matrix optimization problem defined in Lemma D.1 is given by*

$$\mathbf{\Sigma}_{z|x_i} = \mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{U}_i \ \operatorname{diag}\left(\left\{\min\left(1, \frac{\chi_i}{2d_k^i}\right)\right\}_{k=1}^N\right)\mathbf{U}_i'\mathbf{\Sigma}_z^{\frac{1}{2}} \tag{D.21}$$

Proof. Define an instrument matrix

$$\widehat{\mathbf{\Sigma}}_i = \mathbf{U}_i'\mathbf{\Sigma}_z^{-\frac{1}{2}}\mathbf{\Sigma}_{z|x_i}\mathbf{\Sigma}_z^{-\frac{1}{2}}\mathbf{U}_i \succeq 0$$

It's clear that

$$\mathbf{\Sigma}_{z|x_i} = \mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{U}_i\widehat{\mathbf{\Sigma}}_i\mathbf{U}_i'\mathbf{\Sigma}_z^{\frac{1}{2}}. \tag{D.22}$$

Using basic properties of the trace operator,

$$\operatorname{tr}\left(\mathbf{\Omega}_{iz}\mathbf{\Sigma}_{z|x_i}\right) = \operatorname{tr}\left(\mathbf{\Omega}_{iz}\mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{U}_i\widehat{\mathbf{\Sigma}}_i\mathbf{U}_i'\mathbf{\Sigma}_z^{\frac{1}{2}}\right) = \operatorname{tr}\left(\mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{\Omega}_{iz}\mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{U}_i\widehat{\mathbf{\Sigma}}_i\mathbf{U}_i'\right) = \operatorname{tr}\left(\mathbf{U}_i\mathbf{\Omega}_d^i\mathbf{U}_i'\mathbf{U}_i\widehat{\mathbf{\Sigma}}_i\mathbf{U}_i'\right) = \operatorname{tr}\left(\mathbf{\Omega}_d^i\widehat{\mathbf{\Sigma}}_i\right);$$

Similarly, using properties of matrix determinant,

$$\log\det\mathbf{\Sigma}_z - \log\det\mathbf{\Sigma}_{z|x_i} = \log\det\widehat{\mathbf{\Sigma}}_i$$

[34] This result holds for arbitrarily-correlated prior matrix $\mathbf{\Sigma}_z$ and for problems with arbitrary number of optimal decisions such that $\operatorname{rank}(\mathbf{\Omega}_{iz}) > 1$. See Miao et al. (2022) for more details.

The "no-forgeting" constraint $\mathbf{\Sigma}_{z|x_i} \preceq \mathbf{\Sigma}_z$ can also be recasted as

$$\mathbf{\Sigma}_{z|x_i} \preceq \mathbf{\Sigma}_z \Leftrightarrow \mathbf{\Sigma}_z^{\frac{1}{2}} \mathbf{U}_i \widehat{\mathbf{\Sigma}_i} \mathbf{U}_i' \mathbf{\Sigma}_z^{\frac{1}{2}} \preceq \mathbf{\Sigma}_z^{\frac{1}{2}} \mathbf{U}_i \mathbf{U}_i' \mathbf{\Sigma}_z^{\frac{1}{2}} \Leftrightarrow \widehat{\mathbf{\Sigma}_i} \preceq \mathbf{I}$$

where we use the properties of positive-semidefinite matrices. Therefore, we convert the optimization problem in terms of surrogate matrix $\widehat{\mathbf{\Sigma}_i}$,

$$\min_{\widehat{\mathbf{\Sigma}_i}} \operatorname{tr}\left(\mathbf{\Omega}_d^i \widehat{\mathbf{\Sigma}_i}\right) - \frac{1}{2}\chi_i \log\det \widehat{\mathbf{\Sigma}_i}, \tag{D.23}$$

subject to a semidefinite constraint,

$$\widehat{\mathbf{\Sigma}_i} \preceq \mathbf{I}. \tag{D.24}$$

The optimization is a well-defined convex progamming problem, because log det function is concave and (D.24) defines a convex (positive, semidefinite) cone. Recall that $\mathbf{\Omega}_d^i$ is a diagonal matrix of eigenvalues, hence off-diagonal elements in $\widehat{\mathbf{\Sigma}_i}$ does not affect the value of $\operatorname{tr}\left(\mathbf{\Omega}_d^i \widehat{\mathbf{\Sigma}_i}\right)$. Meanwhile, the Hadamard inequality for positive definite matrices implies the determinant of $\widehat{\mathbf{\Sigma}_i}$ is bounded by the product of its diagonal elements,

$$\det \widehat{\mathbf{\Sigma}_i} \leq \prod_{j=1}^{N} \widehat{\mathbf{\Sigma}_i}\left(j, j\right).$$

The equality holds if and only if $\widehat{\mathbf{\Sigma}_i}$ is diagonal. Therefore, it is straightforward to conclude that the optimal solution to (D.23) and (D.24) must be diagonal. This observation allow us to convert the optimization into an equivalent form,

$$\min_{\left\{\widehat{\mathbf{\Sigma}_i}(k,k)\right\}_{k=1}^{N}} \sum_{k=1}^{N} d_k^i \widehat{\mathbf{\Sigma}_i}\left(k, k\right) - \frac{1}{2}\chi_i \sum_{k=1}^{N} \log \widehat{\mathbf{\Sigma}_i}\left(k, k\right)$$

subject to constraints

$$\widehat{\mathbf{\Sigma}_i}\left(k, k\right) \leq 1; \qquad k = 1, 2, ...N$$

By convexity, the Kuhn-Tucker conditions allow us to deliver the unique solution to this problem,

$$\widehat{\mathbf{\Sigma}_i}\left(k, k\right) = \min\left(1, \frac{\chi_i}{2d_k^i}\right); \qquad k = 1, 2, ...N$$

Subtitute this expression into the original solution (D.22) yields Lemma's claim as desired. □

Now given the solution in the preceding lemma, in our rank-one model the solution reduces to

$$\mathbf{\Sigma}_{z|x_i} = \mathbf{\Sigma}_z^{\frac{1}{2}} \mathbf{U}_i \widehat{\mathbf{\Sigma}_i} \mathbf{U}_i' \mathbf{\Sigma}_z^{\frac{1}{2}} = \mathbf{\Sigma}_z^{\frac{1}{2}} \mathbf{U}_i \begin{bmatrix} \min\left(1, \frac{\chi_i}{2d_1^i}\right) & 0 \\ 0 & \mathbf{I}_{N-1} \end{bmatrix} \mathbf{U}_i' \mathbf{\Sigma}_z^{\frac{1}{2}} \tag{D.25}$$

since $d^i_k = 0; \forall k = 2, 3, ...N$. It follows that that when the exogenous marginal cost $\frac{\chi_i}{2d^i_1} \geq 1$ is large enough, $\widehat{\boldsymbol{\Sigma}}_i = I$ and $\boldsymbol{\Sigma}_{z|x_i} = \boldsymbol{\Sigma}_z$. That is, no information is collected (posterior equals prior) when information cost is too high. Using the definition of positive eigenvalue d^i_1, we arrives at the condition for no information acquisition,

$$\chi_i \geq \lambda_i \theta_i \mathbb{V}\left(\mathrm{mc}_{it}\right)$$

as desired. On the other hand, the required threshold condition for processing information is written as

$$\chi_i < 2d^i_1 \tag{D.26}$$

In this case, firms would choose to collect information.[35] By the Bayesian updating formula,

$$\mathrm{SNR}_i = \mathbf{H_i}'\mathbf{V}_i^{-1}\mathbf{H_i} = \boldsymbol{\Sigma}_{z|x_i}^{-1} - \boldsymbol{\Sigma}_z^{-1} = \boldsymbol{\Sigma}_z^{-\frac{1}{2}}\mathbf{U}_i \begin{bmatrix} \left(\frac{2d^i_1}{\chi_i} - 1\right) & 0 \\ 0 & \mathbf{0}_{N-1} \end{bmatrix} \mathbf{U}_i' \boldsymbol{\Sigma}_z^{-\frac{1}{2}}$$

By construction, the signal-to-noise ratio matrix (SNR_i) is rank-one with one strictly positive singular value. Therefore, the optimal signal dimension is 1,

$$\dim\left(\boldsymbol{x}_{it}\right) = \mathrm{rank}\left(\mathrm{SNR}_i\right) = 1$$

by Proposition 1 of Miao et al. (2022) and Tanaka et al. (2017). Next, we partition the unitary matrix $\mathbf{U}_i$ comfortably as $\mathbf{U}_i = \begin{bmatrix} \zeta^i_1 & \zeta^i_2 \end{bmatrix}$ where eigenvector ζ^i_1 corresponds to d^i_1 as we defined before, then the matrix SNR is given by

$$\mathbf{H_i}'\mathbf{V}_i^{-1}\mathbf{H_i} = \boldsymbol{\Sigma}_z^{-\frac{1}{2}}\zeta^i_1\left(\frac{2d^i_1}{\chi_i} - 1\right)(\zeta^i_1)'\,\boldsymbol{\Sigma}_z^{-\frac{1}{2}}$$

Therefore, a natural solution for the univariate optimal signal is

$$\mathbf{V}_i = \frac{\chi_i}{(2d^i_1 - \chi_i)}; \qquad \mathbf{H_i} = (\zeta^i_1)'\,\boldsymbol{\Sigma}_z^{-\frac{1}{2}} = \frac{G_i}{\left\|\boldsymbol{\Sigma}_z^{\frac{1}{2}} G_i'\right\|},$$

Note that the optimal signal structure can also be computed by the singular-value decomposition. In our univariate case, the optimal signal is unique modulo a scalar normalization constant. When $\mathbf{H_i}$ is scaled by a constant b, $\mathbf{V}_i$ is scaled by b^2. Such normalization yields informationally-equivalent signals that produce identical posteriors, and does not affect the sectoral and aggregate equilibrium of our model.[36]

Given this observation and the fact that $\mathrm{mc}_{it} = p^{\Delta}_{it} = R_i^{-\frac{1}{2}}\mathbf{G}_i \boldsymbol{z}_t$, we let $b = R_i^{-\frac{1}{2}}\left\|\boldsymbol{\Sigma}_z^{\frac{1}{2}} G_i'\right\|$. Accordingly, the noise

[35] Of course, d^i_1 is an endogenous objective to be determined in the general equilibrium. In our proof of equilibrium existence in Proposition 3.5 , we provide conditions for χ_i such that (D.26) always holds in general equilibrium.

[36] Normalization does not change the response functions to fundamental shocks, but the responses to idosyncratic noise shocks are scaled by $\frac{1}{b}$. Since idiosyncratic noises wash out at the sectoral and aggregate level, our model equilibrium is unaffected. In Appendix F.1, we provide further evidence.

variance is multiplied by $b^2 = R_i^{-1} \left\| \boldsymbol{\Sigma}_z^{\frac{1}{2}} G_i' \right\|^2 = \mathbb{V}(\mathrm{mc}_{it})$. The normalized optimal signal is given by

$$x_{i\iota t} = \mathrm{mc}_{it} + u_{i\iota t}, \quad u_{i\iota t} \sim \mathbb{N}\left(0, \nu_i^2\right).$$

The endogenous noise variance is given by

$$\nu_i^2 = \frac{\chi_i \mathbb{V}(mc_{it})}{\lambda_i \theta_i \mathbb{V}(mc_{it}) - \chi_i}. \tag{D.27}$$

where we use the fact that $d_1^i = \left\| \boldsymbol{\Sigma}_z^{\frac{1}{2}} \mathbf{G}_i' \right\|^2 = \frac{1}{2}\theta_i \lambda_i \mathbb{V}(mc_{it})$. Note that by construction, the endogenous Gaussian noise $u_{i\iota t}$ is independent of the marginal costs. The noise is also sector-specific in the sense that firms in the same sector face noisy signals drawn from the same distribution. Within-sector information heterogeneity arises ONLY from different realizations of noise $u_{i\iota t}$. The underlying results in the Proposition applies to firms in ALL sectors, $i = 1, 2, ...N$. The proof is now complete. □

D.4 Proof of Corrolary 3.1

Proof. In the preceding analysis, we have derived the optimal pricing decision rules for firm (i, ι) under rational inattention,

$$p_{i\iota t} = \mathbb{E}\left[\mathrm{mc}_{it} \mid x_{i\iota t}\right].$$

Given the optimal signal structure, Bayesian forecasting formula implies that

$$p_{i\iota t} = \mu_i x_{i\iota t} = \mu_i \left(\mathrm{mc}_{it} + u_{i\iota t}\right); \qquad \mu_i = \frac{\mathbb{V}(\mathrm{mc}_{it})}{\mathbb{V}(\mathrm{mc}_{it}) + \nu_i^2} \in [0, 1] \tag{D.28}$$

given the independence between sectoral marginal cost and noises. Substitute the formula of endogenous noise variance into the above expression,

$$\mu_i = 1 - \frac{\chi_i}{\lambda_i \theta_i \mathbb{V}(\mathrm{mc}_{it})}$$

as desired. Note that μ_i is sector specific because all firms in sector i solve an identical information choice problem. Aggregating the individual firms' decisions in sector i, we obtain

$$p_{it} = \mu_i \mathrm{mc}_{it}$$

where idiosyncratic noises wash out $\int u_{i\iota t} d\iota = 0$. The proof is now complete. □

D.5 Proof of Proposition 3.3

Proof. We divide the proof in two parts.

Fixed-capacity case. Using similar approach of Lemma D.1, under fixed-capacity the optimal information choice is given by the following matrix optimization,

$$\begin{aligned}
&\min_{\boldsymbol{\Sigma}_{z|x_i}} \operatorname{tr}\left(\boldsymbol{\Omega}_{iz}\boldsymbol{\Sigma}_{z|x_i}\right)\\
&\text{s.t } 0 \prec \boldsymbol{\Sigma}_{z|x_i} \preceq \boldsymbol{\Sigma}_z,\\
&\log\det\boldsymbol{\Sigma}_z - \log\det\boldsymbol{\Sigma}_{z|x_i} \leq 2\delta_i
\end{aligned}$$

In information theory, the above optimization leads to a classical distortion rate function (DRF),

$$D^*\left(\delta_i\right) = \min_{\boldsymbol{\Sigma}_{z|x_i}} \operatorname{tr}\left(\boldsymbol{\Omega}_{iz}\boldsymbol{\Sigma}_{z|x_i}\right)$$

The problem can solved using the same techniques in Appendix D.3.[37] The optimal distortion is given by

$$D^* = e^{-2\delta_i}||\boldsymbol{\Sigma}_z^{\frac{1}{2}}\mathbf{G}_i'||^2 \tag{D.29}$$

The optimal signal is still one-dimensional with form

$$x_{i\iota t} = \mathrm{mc}_{it} + u_{i\iota t}; \qquad u_{i\iota t} \sim \mathbb{N}\left(0, \nu_i^2\right)$$

The signal noise variance is given by

$$\nu_i^2 = R_i^{-1}\frac{D^*||\boldsymbol{\Sigma}_z^{\frac{1}{2}}\mathbf{G}_i'||^2}{||\boldsymbol{\Sigma}_z^{\frac{1}{2}}\mathbf{G}_i'||^2 - D^*}$$

Therefore, we can compute the signal precision (attention) as

$$\mu_i = \frac{1/\nu_i^2}{1/\nu_i^2 + 1/\mathbb{V}\left(\mathrm{mc}_{it}\right)} = \frac{||\Sigma_0^{\frac{1}{2}}\mathbf{G}_i'||^2 - D^*}{||\Sigma_0^{\frac{1}{2}}\mathbf{G}_i'||^2} \in (0,1)$$

by recalling that $\mathbb{V}\left(\mathrm{mc}_{it}\right) = R_i^{-1}\left\|\Sigma_0^{\frac{1}{2}}\mathbf{G}_i'\right\|^2$. Using (D.29), we obtain,

$$\mu_i = 1 - e^{-2\delta_i}$$

[37] In particular, we solve the following rate distortion function problem,

$$\begin{aligned}
&\delta^*\left(D\right) = \min_{\boldsymbol{\Sigma}_{z|x_i}} \log\det\boldsymbol{\Sigma}_z - \log\det\boldsymbol{\Sigma}_{z|x_i}\\
&\text{s.t } 0 \prec \boldsymbol{\Sigma}_{z|x_i} \preceq \boldsymbol{\Sigma}_z,\\
&\operatorname{tr}\left(\boldsymbol{\Omega}_{iz}\boldsymbol{\Sigma}_{z|x_i}\right) \leq D
\end{aligned}$$

which leads to a rate distortion function (RDF). The two problems are equivalent in the sense that $\delta^*\left(D\right)$ and $D^*\left(\delta_i\right)$ are inverse functions of each other that define a one-to-one, monotone, continous mapping. The solution technique is identical to the elastic attention model, except that the marginal cost of information (Lagrange multiplier) is endogenously determined. The details of solution is available upon request.

as desired.

Exogenous Information Case: For each firm ι in sector i, it observes N independent signals about $\boldsymbol{z}_t$,

$$x_{i\iota jt} = z_{j,t} + u_{i\iota jt}, \quad u_{i\iota jt} \sim \mathbb{N}\left(0, \tau_i \sigma_j^2\right).$$

This implies firm's forecats about sector shocks are given by

$$\mathbb{E}_{i\iota t}\left[z_{jt}|x_{i\iota jt}\right] = \frac{\sigma_j^2}{\sigma_j^2 + \tau_i \sigma_j^2} x_{i\iota jt} = \frac{1}{1+\tau_i} x_{i\iota jt}; \qquad j = 1, 2, ...N \tag{D.30}$$

Therefore, the signal precisions for each sector shock are identical. In Appendix D.2, we show that the pricing rule (D.14) holds for any information structure, with marginal cost function given by (3.3). Therefore, individual firms set price according to

$$p_{i\iota t} = \mathbb{E}_{i\iota t}\left[\mathrm{mc}_{it}|\left\{x_{i\iota jt}\right\}_{j=1}^N\right]$$

Without loss of generality, assume the equilibium function of marginal cost is given by $\mathrm{mc}_{it} = \boldsymbol{m}'\boldsymbol{z}_t = \sum_{j=1}^N m_j z_{jt}$. Then the pricing function follows

$$p_{i\iota t} = \mathbb{E}_{i\iota t}\left[\sum_{j=1}^N m_j z_{jt}|\left\{x_{i\iota jt}\right\}_{j=1}^N\right] = \sum_{j=1}^N m_j \mathbb{E}_{i\iota t}\left[z_{jt}|\left\{x_{i\iota jt}\right\}_{j=1}^N\right] = \sum_{j=1}^N m_j \frac{1}{1+\tau_i} x_{i\iota jt} = \frac{1}{1+\tau_i}\sum_{j=1}^N m_j\left(z_{j,t} + u_{i\iota jt}\right) \tag{D.31}$$

Integrate over all firms in the sector, idiosyncratic noises wash out. Hence,

$$p_{it} = \int \frac{1}{1+\tau_i}\sum_{j=1}^N m_j\left(z_{j,t} + u_{i\iota jt}\right) d\iota = \frac{1}{1+\tau_i}\sum_{j=1}^N m_j z_{jt} = \frac{1}{1+\tau_i}\mathrm{mc}_{it} \tag{D.32}$$

where the last equality follows from our initial conjecture. The proof is therefore complete. □

D.6 Proof of Proposition 3.4

Proof. In our previous derivations, we conjecture that the equilibrium sectoral price follows $\mathbf{p}_t = \phi \boldsymbol{z}_t$. The marginal cost , or the ideal price, is given by

$$\mathbf{mc}_t = -\boldsymbol{z}_t + \boldsymbol{\alpha} w_t + \mathbf{A}\mathbf{p}_t,$$

which generalizes (3.3). Meanwhile, Corrolary 3.1 indicates that equilibrium price vector is given by

$$\mathbf{p}_t = \boldsymbol{\mu}\mathbf{mc}_t$$

where $\boldsymbol{\mu} = \text{diag}\left(\{\mu_i\}_{i=1}^N\right)$. Combining these two equation leads to

$$\mathbf{p}_t = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1} \boldsymbol{\mu} (-\mathbf{I} + \alpha\boldsymbol{\kappa}) \boldsymbol{z}_t; \qquad \phi = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1} \boldsymbol{\mu} (-\mathbf{I} + \alpha\boldsymbol{\kappa}) \tag{D.33}$$

Therefore, the marginal cost can be solved as

$$\mathbf{mc}_t = \left(\mathbf{I} + \mathbf{A}(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1} \boldsymbol{\mu}\right) (-\mathbf{I} + \alpha\boldsymbol{\kappa}) \boldsymbol{z}_t$$

Next, we present a technical Lemma using the theory of nonnegative matrices. This result will be applied repeatedly in subsequent analysis.

Lemma D.3. *For any vector $\boldsymbol{\mu}^v \in [0,1]^N$, the spectral radius of the non-negative matrix $\boldsymbol{\mu}\mathbf{A}$ is strictly smaller than 1, $\rho(\boldsymbol{\mu}\mathbf{A}) < 1$. Therefore, $\mathbf{I} - \boldsymbol{\mu}\mathbf{A}$ is a M matrix.*

Proof. By definition,

$$\boldsymbol{\mu}\mathbf{A} = \begin{bmatrix} a_{11}\mu_1 & a_{12}\mu_1 & \dots & a_{1N}\mu_1 \\ a_{21}\mu_2 & a_{22}\mu_2 & \dots & a_{2N}\mu_2 \\ \vdots & \vdots & \ddots & \dots \\ a_{N1}\mu_N & a_{N2}\mu_N & \dots & a_{NN}\mu_N \end{bmatrix}$$

is nonnegative. Therefore, we employ the following Theorem in matrix analysis,

Theorem (Gershgorin Circle Theorem). *For an $N \times N$ matrix $\mathbb{L}$, each eigenvalue z of $\mathbb{L}$ satisfy the following conditions,*

$$|z - l_{ii}| \leq \sum_{j=1;j\neq i}^{N} |l_{ii}|, \quad i \in \{1, 2, \cdots, N\}.$$

where l denotes element of $\mathbb{L}$.

By the Gershgorin circle theorem and the Perron-Frobenius Theorem (Berman and Plemmons (1994), Chapter 2), the absolute value of maximal eigenvalue of $\boldsymbol{\mu}\mathbf{A}$, or its spectral radius $\rho(\boldsymbol{\mu}\mathbf{A})$, is bounded above,[38]

$$\rho(\boldsymbol{\mu}\mathbf{A}) \leq \max_i \left\{ \sum_{j=1}^{N} a_{ij}\mu_j : i = 1, 2, \ldots N \right\} \leq \max_i \left\{ \sum_{j=1}^{N} a_{ij} : i = 1, 2, \ldots N \right\} = \|\mathbf{A}\|_\infty < 1$$

[38] If $\boldsymbol{\mu}\mathbf{A} >> 0$ is strictly positive, then the Perron-Frobenius Theorem implies that there exists a unique positive and maximal (dominant) eigenvalue that is equals to its spectral radius.

where constant $\|\mathbf{A}\|_\infty \in (0,1)$ is the matrix operator norm induced by the maximal row-sum of $\mathbf{A}$.[39]

$$\|\mathbf{A}\|_\infty = \max_i \left\{ \sum_{j=1}^{N} |a_{ij}| : i = 1, 2, \ldots N \right\} = \max_i \left\{ \sum_{j=1}^{N} a_{ij} : i = 1, 2, \ldots N \right\} \tag{D.34}$$

By definition in Plemmons (1977), $(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})$ is a M matrix, such that $(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}$ exists and is nonnegative. In other words, the matrix inverse admits series expansion as

$$(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1} = \sum_{n=0}^{\infty} (\boldsymbol{\mu}\mathbf{A})^n = \mathbf{I} + (\boldsymbol{\mu}\mathbf{A}) + (\boldsymbol{\mu}\mathbf{A})^2 + \ldots. > 0$$

□

Given the matrix expansion, we have

$$\mathbf{I} + \mathbf{A}(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu} = \mathbf{I} + \mathbf{A}\left(\mathbf{I} + (\boldsymbol{\mu}\mathbf{A}) + (\boldsymbol{\mu}\mathbf{A})^2 + \ldots.\right)\boldsymbol{\mu} = \mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + \ldots. = (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} > 0 \tag{D.35}$$

where the matrix $(\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}$ is also a M matrix with spectral radius bounded by 1. We also have an important matrix identity

$$\boldsymbol{\mu}(\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}, \tag{D.36}$$

which will be used repeatedly in subsequent analysis. Note that above results DOES NOT rely on the invertibility of $\boldsymbol{\mu}$. In addition, the volatility of marginal cost is now given by

$$\mathbb{V}(mc_{i,t}) = \left\| \mathbf{e}_i(\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}(-\mathbf{I} + \boldsymbol{\alpha\kappa})\boldsymbol{\Sigma}_z^{\frac{1}{2}} \right\|^2,$$

To summarize, the fixed-point of nominal rigidities (sectoral attentions) is given by (3.3) while influence matrix ϕ is given by (D.33), as desired. The proof is now complete. □

D.7 Proof of Corrolary 3.2

Proof. First, we derive consumption deviations. In rationally inattentive equilibrium, the nominal GDP implied by household's budget constraint follow

$$P_t^f C_t = W_t L_t + \sum_i P_{it} Y_{it} (1 - \mathcal{E}_{it}) = W_t L_t \left(1 - \sum_{i=1}^{N} \Lambda_{it} (1 - \mathcal{E}_{it})\right)^{-1} \tag{D.37}$$

where sectoral markup deviations are defined as $\mathcal{E}_{it} \equiv \frac{\mathrm{MC}_{it}}{P_{it}} \int \left(\frac{P_{it\iota}}{P_{it}}\right)^{-\theta_i} d\iota$ and sectoral Domar weights are given by $\Lambda_{it} = \frac{P_{it} Y_{it}}{P_t^f C_t}$ under rational inattention. Standard log-linearization leads to $1 - \sum_{i=1}^{N} \Lambda_{it}(1 - \mathcal{E}_{it}) \approx 1 + \sum_{i=1}^{N} \lambda_i \varepsilon_{it}$.

[39] Here we adopt Horn and Johnson (2012)'s notation on matrix norms.

Meanwhile, log-linearized sectoral markups are given by

$$\varepsilon_{it} = mc_{it} - p_{it} - \theta_i \int \left(p_{\iota t} - p_{it}\right) = mc_{it} - p_{it}$$

We immediately have

$$\sum_{i=1}^{N} \lambda_i \varepsilon_{it} = \sum_{i=1}^{N} \lambda_i \left(-z_{it} + \alpha_i w_t + \sum_{j=1}^{N} a_{ij} p_{jt} - p_{it}\right) = -p_t^f + \sum_{i=1}^{N} \lambda_i \left(-z_{it} + \alpha_i w_t\right) = w_t - p_t^f - \sum_{i=1}^{N} \lambda_i z_{it} \tag{D.38}$$

where we use the fact that $\sum_{i=1}^{N} \lambda_i \alpha_i = 1$. Under rational inattention, the log-linearized budget constraint reads

$$c_t = w_t - p_t^f + \ell_t - \sum_{i=1}^{N} \lambda_i \varepsilon_{it}$$

subustitue (D.38) and the Euler equation into this equation, we get

$$c_t = \frac{\eta}{1+\gamma\eta} \left(w_t - p_t^f\right) + \frac{1}{1+\gamma\eta} \sum_i \lambda_i z_{it}$$

In the proof of Proposition 3.1, we show that the equilibrium consumption is given by

$$c_t^* = \frac{1+\eta}{1+\gamma\eta} \left(w_t^* - p_t^{f,*}\right) = \frac{1+\eta}{1+\gamma\eta} \boldsymbol{\lambda}' \boldsymbol{z}_t$$

In FI flexible-price economy, without monetary policy only the relative price $\left(w_t^* - p_t^{f,*}\right)$ is determiate. Introducing CIA constraint pins down wage and final good price index separately; however, money neutrality implies that the relative price and the consumption (and labor) are independent of monetary policy. Without loss of generality, we let $w_t^* = w_t$, then the consumption deviation becomes,

$$c_t - c_t^* = -\frac{\eta}{1+\gamma\eta} \left(p_t^f - p_t^{f,*}\right) = -\frac{\eta}{1+\gamma\eta} \sum_{i=1}^{N} \beta_i e_{it}$$

as desired. Note the consumption deviation is independent of the choice of wage function in the FI economy. We will expolit this property in the subsequent welfare analysis. Finally, we combine (D.9) and(D.33) and simplify the algebra, we obtain that

$$\begin{aligned}
\mathbf{e}_t &= \Big((\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1} \boldsymbol{\mu}(\mathbf{I} - \mathbf{A})(-\mathbf{L} + \mathbf{1}\boldsymbol{\kappa}) - (\mathbf{I} - \mathbf{A})^{-1}(-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa})\Big) \boldsymbol{z}_t \\
&= \Big((\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}) - (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1} \boldsymbol{\mu}(\mathbf{I} - \mathbf{A})(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa})\Big) \boldsymbol{z}_t \\
&= \Big(\mathbf{I} - (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1} \boldsymbol{\mu}(\mathbf{I} - \mathbf{A})\Big)(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}) \boldsymbol{z}_t
\end{aligned}$$

$$
\begin{aligned}
&= \left((\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I}-\boldsymbol{\mu}\mathbf{A}-\boldsymbol{\mu}(\mathbf{I}-\mathbf{A}))\right)(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})\boldsymbol{z}_t \\
&= \left((\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I}-\boldsymbol{\mu})\right)(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})\boldsymbol{z}_t \\
&= \mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{z}_t
\end{aligned}
$$

The proof is now complete. □

D.8 Proof of Proposition 3.5

Proof. I present the proofs of this proposition in two parts. The first part proves the equilibrium existence. The second part establishes equilibrium uniqueness.

Part I: Proof of Existence.

Recall from (3.6), the price rigidities $\{\mu_i\}_{i=1}^N$ solve the following fixed-point problem

$$
\mu_i = 1 - \frac{\chi_i}{\theta_i \lambda_i \mathbb{V}(\mathrm{mc}_{it})}, \qquad \text{and} \qquad \mathbb{V}(\mathrm{mc}_{it}) = \left\| \mathbf{e}_i (\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1}(-\mathbf{I}+\boldsymbol{\alpha\kappa})\boldsymbol{\Sigma}_z^{\frac{1}{2}} \right\|^2 .
$$

Given the N-dimensional column vector of interest, $\boldsymbol{\mu}^v = \left(\mu_1, \mu_2, ...\mu_N\right)' \in \mathbb{R}^N$, we define an N-dimensional vector-valued function $\mathcal{T}\left(\boldsymbol{\mu}^v\right) : [0,1]^N \longmapsto \mathbb{R}^N$ as

$$
\mathcal{T}\left(\boldsymbol{\mu}^v\right) = \left(\mathcal{T}_1\left(\boldsymbol{\mu}\right), \mathcal{T}_2\left(\boldsymbol{\mu}\right), ...\mathcal{T}_N\left(\boldsymbol{\mu}\right)\right)'; \tag{D.39}
$$

$$
\mathcal{T}_i\left(\boldsymbol{\mu}^v\right) = 1 - \frac{\chi_i}{\theta_i \lambda_i \left\| \mathbf{e}_i (\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1}(-\mathbf{I}+\boldsymbol{\alpha\kappa})\boldsymbol{\Sigma}_z^{\frac{1}{2}} \right\|^2}. \tag{D.40}
$$

Then the fixed-point system can be written as

$$
\mathcal{T}\left(\boldsymbol{\mu}^v\right) = \boldsymbol{\mu}^v.
$$

By inspection, $\mathcal{T}\left(\boldsymbol{\mu}\right)$ is continuous on $[0,1]^N$, and the Cartesian cube $[0,1]^N \subseteq \mathbb{R}^N$ is compact and convex. Therefore, we invoke the Brouwer fixed-point theorem in the N-dimensional Euclidean (Banach) space $\mathbb{R}^N$. We only need to show that $\mathcal{T}\left(\boldsymbol{\mu}\right)$ defines a self-map from $[0,1]^N$ onto itself.

Using Lemma D.3, it is easy to see that $(\mathbf{I}-\mathbf{A}\boldsymbol{\mu})$ is a M matrix, such that $(\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1}$ exists and is nonnegative. In other words, the inverse admits the series expansion as

$$
(\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1} = \sum_{n=0}^{\infty} (\mathbf{A}\boldsymbol{\mu})^n = \mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ... > 0.
$$

Now define the attention-distorted Leontief matrix,

$$
\boldsymbol{\Delta}_\mu = (\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1} > 0, \tag{D.41}
$$

and the wage-related coefficient matrix $\mathbf{\Gamma} = (-\mathbf{I} + \boldsymbol{\alpha\kappa})$. Then follows a proposition.

Proposition D.1. *If the monetary policy (wage) rule satisfy*

$$\kappa_i < 1; \qquad \forall i = 1, 2, ...N,$$

then each sector's ideal price (marginal cost) volatility has a positive lower bound,

$$\mathbb{V}(mc_{it}) = \left\| \mathbf{e}_i (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} (-\mathbf{I} + \boldsymbol{\alpha\kappa}) \mathbf{\Sigma}_z^{\frac{1}{2}} \right\|^2 \geq \sigma_i^2 (1 - \varkappa_i \kappa_i)^2 > 0,$$

where the auxiliary parameters $\{\varkappa_i\}$ is defined as

$$\varkappa_i = \begin{cases} 1, & 0 \leq \kappa_i < 1 \\ \alpha_i, & \kappa_i < 0 \end{cases}.$$

Proof. By definition of the Euclidean norm,[40]

$$\mathbb{V}(\mathrm{mc}_{it}) = \left\| \mathbf{e}_i (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} (-\mathbf{I} + \boldsymbol{\alpha\kappa}) \mathbf{\Sigma}_z^{\frac{1}{2}} \right\|^2 = \left\| \mathbf{e}_i \mathbf{\Delta}_\mu \mathbf{\Gamma} \mathbf{\Sigma}_z^{\frac{1}{2}} \right\|^2 \geq \left\| \left[\mathbf{\Delta}_\mu \mathbf{\Gamma} \right]_{ii} \mathbf{e}_i \mathbf{\Sigma}_z^{\frac{1}{2}} \right\|^2 \tag{D.42}$$

where the scalar operator $\left[\,\cdot\,\right]_i$ or $\left[\,\cdot\,\right]_{ii}$ denotes the i-th or ii-th element of the vector or matrix, and we use the fact that $\mathbf{\Sigma_z} > 0$ is diagonal under our baseline assumption. Next, we observe that

$$\begin{aligned} \mathbf{\Delta}_\mu \mathbf{\Gamma} &= (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} (-\mathbf{I} + \boldsymbol{\alpha\kappa}) \\ &= \left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ... \right] (-\mathbf{I} + \boldsymbol{\alpha\kappa}) \\ &= - \left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ... \right] + \left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ... \right] (\boldsymbol{\alpha\kappa}). \end{aligned} \tag{D.43}$$

By definition of $\mathbf{A}$ and $\boldsymbol{\mu}$,

$$0 \leq \left[(\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ... \right] \leq \left[(\mathbf{A}) + (\mathbf{A})^2 + ... \right].$$

It is immediate that (D.43) is bounded by

$$\mathbf{\Delta}_\mu \mathbf{\Gamma} \leq -\mathbf{I} + \left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ... \right] (\boldsymbol{\alpha\kappa}).$$

Notice that the ith row of matrix $\left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ... \right] (\boldsymbol{\alpha\kappa})$ is $\left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ... \right] \boldsymbol{\alpha}\kappa_i$, for $i = 1, 2, ...N$. On one hand, when $\kappa_i < 0$,

$$\left[\mathbf{\Delta}_\mu \mathbf{\Gamma} \right]_{ii} \leq \left[-\mathbf{I} + \left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ... \right] (\boldsymbol{\alpha\kappa}) \right]_{ii} = -1 + \left[\left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ... \right] \boldsymbol{\alpha}\kappa_i \right]_i \leq -1 + \alpha_i \kappa_i < 0, \tag{D.44}$$

[40] The volatility bound we establish here is obviously non-unique and can be futher relaxed, but the current condition suffices for our purpose.

On the other hand, when $0 \le \kappa_i < 1$, the column vector $\left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ...\right] \boldsymbol{\alpha}\kappa_i$ is bounded as

$$\left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ...\right] \boldsymbol{\alpha}\kappa_i \le \left[\mathbf{I} + (\mathbf{A}) + (\mathbf{A})^2 + ...\right] \boldsymbol{\alpha}\kappa_i = (\mathbf{I} - \mathbf{A})^{-1} \boldsymbol{\alpha}\kappa_i = \mathbf{1}\kappa_i,$$

where I use the matrix property of $(\mathbf{I} - \mathbf{A})^{-1} \boldsymbol{\alpha} = \mathbf{1} = (1, 1, ...1)'$. In this case,

$$\left[\boldsymbol{\Delta}_\mu \boldsymbol{\Gamma}\right]_{ii} \le \left[- \mathbf{I} + \left[\mathbf{I} + (\mathbf{A}\boldsymbol{\mu}) + (\mathbf{A}\boldsymbol{\mu})^2 + ...\right] (\boldsymbol{\alpha}\boldsymbol{\kappa}) \right]_{ii} \le -1 + [\kappa_i \mathbf{1}]_i = -1 + \kappa_i < 0, \tag{D.45}$$

where the last inequality follows from the assumption that $\kappa_i < 1$. Hence, by (D.44) and (D.45), all diagonal elements of matrix $\boldsymbol{\Delta}_\mu \boldsymbol{\Gamma}$ are negative under the assumption of $\kappa_i < 1$, that is,

$$- \left[\boldsymbol{\Delta}_\mu \boldsymbol{\Gamma}\right]_{ii} > 0; \qquad \forall i = 1, 2, ...N.$$

Therefore, using (D.42), we conclude that

$$\mathbb{V}(\mathrm{mc}_{it}) = \left\| \mathbf{e}_i \boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z^{\frac{1}{2}} \right\|^2 \ge \left\| \left[\boldsymbol{\Delta}_\mu \boldsymbol{\Gamma}\right]_{ii} \mathbf{e}_i \boldsymbol{\Sigma}_z^{\frac{1}{2}} \right\|^2 \ge \begin{cases} \sigma_i^2 (1 - \kappa_i)^2, & 0 \le \kappa_i < 1 \\ \sigma_i^2 (1 - \alpha_i \kappa_i)^2, & \kappa_i < 0 \end{cases}$$

The proof is complete. □

By Proposition D.1, when $\boldsymbol{\kappa}$ satisfies $\kappa_i < 1, \forall i$,

$$0 < \frac{\chi_i}{\theta_i \lambda_i \mathbb{V}(\mathrm{mc}_{it})} \le \frac{\chi_i}{\theta_i \lambda_i \sigma_i^2 (1 - \varkappa_i \kappa_i)^2} < 1, \tag{D.46}$$

where the last inequality follow directly from our sufficiency condition in Proposition 3.5. The first inequality follows that $(\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}$ exists and is finite, the sectoral volatilities are finite, and that all sectors face non-zero information cost, i.e. $\chi_i > 0, \forall i$.

It is then immediate that for all sector $i = 1, 2, ...N$ and all vectors $\boldsymbol{\mu}^v \in [0, 1]^N$,

$$\mathcal{T}_i(\boldsymbol{\mu}^v) = 1 - \frac{\chi_i}{\theta_i \lambda_i \left\| \mathbf{e}_i (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} (-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa}) \boldsymbol{\Sigma}_z^{\frac{1}{2}} \right\|^2} \in (0, 1), \qquad \forall i = 1, 2, ...N; \tag{D.47}$$

therefore, the N-dimensional vector-valued function $\mathcal{T}(\boldsymbol{\mu}^v) : [0, 1]^N \longmapsto [0, 1]^N$ is a continuous self-map. Then by the Brouwer fixed-point theorem (Hutson et al. (2005), Theorem 8.1.1), the equilibrium system in condition (3.6) has a (interior) fixed-point.

Part II: Proof of Uniqueness.

The key novel result in Proposition 3.5 is the global uniqueness of general equilibrium. In models with endogenous information frictions (e.g. RI or learning from endogenous price), unique equilibrium is hard to prove and characterize, if not impossible (Adams (2019)). Due to the nonlinearity of equilibrium operator arosen from endogenous information, the Contraction Mapping Theorem cannot be applied. Instead, we employ the following extension of the Schauder's fixed-point theorem by Kellogg (1976):

Kellogg's Fixed Point Theorem. *Let X be a real Banach space with a bounded convex open subset $D \subseteq X$. Let $\bar{D}$ be its closure and let $F : \bar{D} \to \bar{D}$ be a compact continuous map which is continuously Fréchet differentiable on D. Suppose that (a) for each $x \in D$, 1 is not an eigenvalue of Fréchet derivative $F'(x)$, and (b) for each $x \in \partial D$ on boundary, $x \neq F(x)$. Then F has a unique fixed point.*

Kellogg's Theorem applies to infinite-dimensional Banach spaces. For discussions of this theorem, see Talman (1978) and Smith and Stuart (1980). As note in Kellogg's original paper, the compactness hypothesis can be dropped in finite-dimensional problems. So in our proof, we will omit the compactness condition.

We proceed with the proof of uniqueness in three steps.

Step 1: Compute the Fréchet derivative. Suppose now $\boldsymbol{\mu}^v \in [0,1]^N$. In our finite-dimensional problem, the Fréchet derivative of the vector-valued function $\mathcal{T}(\boldsymbol{\mu}^v)$ is the $N \times N$ Jacobian matrix,

$$\mathcal{T}'(\boldsymbol{\mu}^v) \equiv \frac{\partial \mathcal{T}}{\partial \boldsymbol{\mu}^v} = \mathcal{T}_{\boldsymbol{\mu}^v} = \begin{bmatrix} \frac{\partial \mathcal{T}_1}{\partial \mu_1} & \frac{\partial \mathcal{T}_1}{\partial \mu_2} & \cdots & \frac{\partial \mathcal{T}_1}{\partial \mu_N} \\ \frac{\partial \mathcal{T}_2}{\partial \mu_1} & \frac{\partial \mathcal{T}_2}{\partial \mu_2} & \cdots & \frac{\partial \mathcal{T}_2}{\partial \mu_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mathcal{T}_N}{\partial \mu_1} & \frac{\partial \mathcal{T}_N}{\partial \mu_2} & \cdots & \frac{\partial \mathcal{T}_N}{\partial \mu_N} \end{bmatrix}. \tag{D.48}$$

On the other hand, we define a matrix of derivatives for the ith component of $\mathcal{T}$ with respect to a $N \times N$ **symmetric** matrix of variable, $\boldsymbol{\mu}$,

$$\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}} = \begin{bmatrix} \frac{\partial \mathcal{T}_i}{\partial \mu_{11}} & \frac{\partial \mathcal{T}_i}{\partial \mu_{12}} & \cdots & \frac{\partial \mathcal{T}_i}{\partial \mu_{1N}} \\ \frac{\partial \mathcal{T}_i}{\partial \mu_{21}} & \frac{\partial \mathcal{T}_i}{\partial \mu_{22}} & \cdots & \frac{\partial \mathcal{T}_i}{\partial \mu_{2N}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mathcal{T}_i}{\partial \mu_{N1}} & \frac{\partial \mathcal{T}_i}{\partial \mu_{N2}} & \cdots & \frac{\partial \mathcal{T}_i}{\partial \mu_{NN}} \end{bmatrix}; \qquad i = 1, 2, ...N$$

here the diagonal elements $\mu_{11}, \mu_{22}, ...\mu_{NN}$ coincide with $\boldsymbol{\mu}^v$. The following definition establishes the relation between the Jacobian matrix and the matrix-valued derivatives.

Definition 3. *The Jacobian matrix $\mathcal{T}_{\boldsymbol{\mu}^v}$ is obtained by extracting the diagonal elements of $\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}$ for each sector, arranged in rows.*

$$\mathcal{T}_{\boldsymbol{\mu}^v} = \frac{\partial \mathcal{T}(\boldsymbol{\mu}^v)}{\partial \boldsymbol{\mu}^v} = \left[\mathrm{diag}\left(\frac{\partial \mathcal{T}_1(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right) \quad \mathrm{diag}\left(\frac{\partial \mathcal{T}_2(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right) \quad \cdots \quad \mathrm{diag}\left(\frac{\partial \mathcal{T}_N(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right)\right]'. \tag{D.49}$$

Relation (D.49) holds either we treat $\boldsymbol{\mu}$ as a general $N \times N$ symmetric matrix of variables or as a diagonal matrix with off-diagonal restrictions of constant 0. Therefore, in the first step we treat $\boldsymbol{\mu}$ as a general $N \times N$ variable matrix, and exploit its diagonal (symmetric) property whenever is necessary.

It suffices to calculate the matrix of derivatives of the scalar function, $\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}$. Recall that $\mathcal{T}_i(\boldsymbol{\mu}) = 1 - \frac{\chi_i}{\theta_i \lambda_i \mathbb{V}(\mathrm{mc}_{it})}$, using the chain rule for matrix derivatives (Magnus and Neudecker (2007)),

$$\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}} = \frac{\chi_i}{\theta_i \lambda_i \mathbb{V}(\mathrm{mc}_{it})^2} \frac{\partial \mathbb{V}(\mathrm{mc}_{it})}{\partial \boldsymbol{\mu}}, \tag{D.50}$$

where the volatility of marginal cost admits matrix representation,

$$\mathbb{V}(\mathrm{mc}_{it}) = \left\| \mathbf{e}_i (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} (-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa}) \boldsymbol{\Sigma}_z^{\frac{1}{2}} \right\|^2 = \left\| \mathbf{e}_i \boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z^{\frac{1}{2}} \right\|^2 = \mathbf{e}_i \boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z \boldsymbol{\Gamma}' \boldsymbol{\Delta}_\mu' \mathbf{e}_i'.$$

Recall that $\boldsymbol{\Delta}_\mu = (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}$ from (D.41). To compute the matrix derivative $\frac{\partial \mathbb{V}(\mathrm{mc}_{it})}{\partial \boldsymbol{\mu}}$, we start with matrix differentials and present 3 sets of basic results using matrix calculus.

Lemma D.4. *Suppose g is a scalar function of $N \times N$ matrix of variable $\boldsymbol{\mu}$, its differential follows,*

$$dg(\boldsymbol{\mu}) = \mathrm{tr}\left(\frac{\partial g(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}'} d\boldsymbol{\mu} \right). \tag{D.51}$$

Proof. By definition,

$$\begin{aligned} dg(\boldsymbol{\mu}) &= \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial g}{\partial \mu_{ij}} d\mu_{ij} \\ &= \mathrm{tr}\left(\begin{bmatrix} \frac{\partial g}{\partial \mu_{11}} & \frac{\partial g}{\partial \mu_{21}} & \cdots & \frac{\partial g}{\partial \mu_{N1}} \\ \frac{\partial g}{\partial \mu_{12}} & \frac{\partial g}{\partial \mu_{22}} & \cdots & \frac{\partial g}{\partial \mu_{N2}} \\ & \cdots & & \\ \frac{\partial g}{\partial \mu_{1N}} & \frac{\partial g}{\partial \mu_{2N}} & \cdots & \frac{\partial g}{\partial \mu_{NN}} \end{bmatrix} \begin{bmatrix} d\mu_{11} & d\mu_{12} & \cdots & d\mu_{1N} \\ d\mu_{21} & d\mu_{22} & \cdots & d\mu_{2N} \\ & & \cdots & \\ d\mu_{N1} & d\mu_{N2} & \cdots & d\mu_{NN} \end{bmatrix} \right) \\ &= \mathrm{tr}\left(\frac{\partial g(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}'} d\boldsymbol{\mu} \right), \end{aligned}$$

where $\mathrm{tr}(\cdot)$ is the trace operator. ◻

The trace operator is linear and has basic properties

$$\begin{aligned} \mathrm{tr}(\boldsymbol{X}\boldsymbol{Y}) &= \mathrm{tr}(\boldsymbol{Y}\boldsymbol{X}); \\ \mathrm{tr}(\boldsymbol{X}) &= \mathrm{tr}(\boldsymbol{X}'), \end{aligned} \tag{D.52}$$

for any $N \times N$ matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$.

Lemma D.5. *For matrix of variables $\boldsymbol{X}$, $\boldsymbol{Y}$, and constant matrices $\mathbf{B}$ and $\mathbf{C}$ of conformable dimensions, we have following differential results,*

$$d(\mathbf{B}\boldsymbol{X}\mathbf{C}) = \mathbf{B}d(\boldsymbol{X})\mathbf{C}; \qquad d\left(\boldsymbol{X}^{-1}\right) = -\boldsymbol{X}^{-1} d(\boldsymbol{X}) \boldsymbol{X}^{-1};$$

$$d(\mathrm{tr}(\boldsymbol{X})) = \mathrm{tr}(d(\boldsymbol{X})); \qquad d(\boldsymbol{X}\boldsymbol{Y}) = d(\boldsymbol{X})\boldsymbol{Y} + \boldsymbol{X}d(\boldsymbol{Y}).$$

Proof. See Magnus and Neudecker (2007), p.160, p.208; Vetter (1973), p.357 □

Lemma D.6. *Given constant N- dimensional column vectors* $\mathbf{b}$ *and* $\mathbf{c}$, *a constant $N \times N$ matrix* $\mathbf{B}$, *and a matrix of variables* $\boldsymbol{X}$, *the scalar function* $\mathbf{b}'\boldsymbol{X}\mathbf{B}\boldsymbol{X}'\mathbf{c}$ *of variable matrix* $\boldsymbol{X}$ *admits differential*

$$d\left(\mathbf{b}'\boldsymbol{X}\mathbf{B}\boldsymbol{X}'\mathbf{c}\right) = \mathrm{tr}\left[\mathbf{b}'d(\boldsymbol{X})\mathbf{B}\boldsymbol{X}'\mathbf{c} + \mathbf{c}'d(\boldsymbol{X})\mathbf{B}'\boldsymbol{X}'\mathbf{b}\right]. \tag{D.53}$$

Proof. Using formulas from the last lemma and the the basic properties of trace operator (linearity and (D.52)), we deduce that

$$\begin{aligned}
d\left(\mathbf{b}'\boldsymbol{X}\mathbf{B}\boldsymbol{X}'\mathbf{c}\right) &= d\left(\mathrm{tr}\left(\mathbf{b}'\boldsymbol{X}\mathbf{B}\boldsymbol{X}'\mathbf{c}\right)\right) \\
&= \mathrm{tr}\left(d\left(\mathbf{b}'\boldsymbol{X}\mathbf{B}\boldsymbol{X}'\mathbf{c}\right)\right) \\
&= \mathrm{tr}\left(\mathbf{b}'d\left(\boldsymbol{X}\mathbf{B}\boldsymbol{X}'\right)\mathbf{c}\right) \\
&= \mathrm{tr}\left[\mathbf{b}'\left(d(\boldsymbol{X})\mathbf{B}\boldsymbol{X}' + \boldsymbol{X}\mathbf{B}d\left(\boldsymbol{X}'\right)\right)\mathbf{c}\right] \\
&= \mathrm{tr}\left[\mathbf{b}'d(\boldsymbol{X})\mathbf{B}\boldsymbol{X}'\mathbf{c}\right] + \mathrm{tr}\left[\mathbf{b}'\boldsymbol{X}\mathbf{B}d\left(\boldsymbol{X}'\right)\mathbf{c}\right] \\
&= \mathrm{tr}\left[\mathbf{b}'d(\boldsymbol{X})\mathbf{B}\boldsymbol{X}'\mathbf{c}\right] + \mathrm{tr}\left[\mathbf{c}'d(\boldsymbol{X})\mathbf{B}'\boldsymbol{X}'\mathbf{b}\right] \\
&= \mathrm{tr}\left[\mathbf{b}'d(\boldsymbol{X})\mathbf{B}\boldsymbol{X}'\mathbf{c} + \mathbf{c}'d(\boldsymbol{X})\mathbf{B}'\boldsymbol{X}'\mathbf{b}\right].
\end{aligned}$$

□

Combining results from Lemma D.4 - D.6, we compute the differential of modified Leontief inverse $\mathbf{\Delta}_\mu = (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}$ as

$$d\left(\mathbf{\Delta}_\mu\right) = -\mathbf{\Delta}_\mu d\left(\mathbf{\Delta}_\mu^{-1}\right)\mathbf{\Delta}_\mu = \mathbf{\Delta}_\mu\mathbf{A}d(\boldsymbol{\mu})\mathbf{\Delta}_\mu;$$

Similarly,

$$d\left(\mathbf{\Delta}_\mu'\right) = \left(d\left(\mathbf{\Delta}_\mu\right)\right)' = \mathbf{\Delta}_\mu'\left(d(\boldsymbol{\mu})\right)'\mathbf{A}'\mathbf{\Delta}_\mu' = \mathbf{\Delta}_\mu'd(\boldsymbol{\mu})\mathbf{A}'\mathbf{\Delta}_\mu',$$

where we exploit the diagonal (symmetric) property of $\boldsymbol{\mu}$ in the last equality. Next, we derive the differential for the volatility of marginal cost,

Lemma D.7. *The volatility of marginal cost,* $\mathbb{V}(mc_{it})$ *admits differential,*

$$d\left[\mathbb{V}(mc_{it})\right] = 2\,\mathrm{tr}\left[\mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{A}d(\boldsymbol{\mu})\right]. \tag{D.54}$$

Proof. Combining results from Lemma D.4-D.6,

$$\begin{aligned}
d\left[\mathbb{V}(\mathrm{mc}_{it})\right] &= d\left(\mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i'\right)\\
&= \mathrm{tr}\left[\mathbf{e}_i\left(d(\mathbf{\Delta}_\mu)\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu' + \mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'d(\mathbf{\Delta}_\mu')\right)\mathbf{e}_i'\right]\\
&= \mathrm{tr}\left[\mathbf{e}_i d(\mathbf{\Delta}_\mu)\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i' + \mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'d(\mathbf{\Delta}_\mu')\mathbf{e}_i'\right]\\
&= 2\,\mathrm{tr}\left[\mathbf{e}_i d(\mathbf{\Delta}_\mu)\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i'\right]\\
&= 2\,\mathrm{tr}\left[\mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{A}d(\boldsymbol{\mu})\mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i'\right]\\
&= 2\,\mathrm{tr}\left[\mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{A}d(\boldsymbol{\mu})\right],
\end{aligned}$$

where I use the rotational property of the trace operator. □

Note that by equation (D.51) in Lemma D.4 and equation (D.54) in Lemma D.7,

$$d\left[\mathbb{V}(\mathrm{mc}_{it})\right] = \mathrm{tr}\left(\frac{\partial\mathbb{V}(\mathrm{mc}_{it})}{\partial\boldsymbol{\mu}'}d\boldsymbol{\mu}\right) = \mathrm{tr}\left(\frac{\partial\mathbb{V}(\mathrm{mc}_{it})}{\partial\boldsymbol{\mu}}d\boldsymbol{\mu}\right) = 2\,\mathrm{tr}\left[\mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{A}d(\boldsymbol{\mu})\right], \tag{D.55}$$

because $\boldsymbol{\mu} = \boldsymbol{\mu}'$. Since $d(\boldsymbol{\mu})$ is a matrix of differentials for variable $\boldsymbol{\mu}$, in order for (D.55) to hold, we must have

$$\frac{\partial\mathbb{V}(\mathrm{mc}_{it})}{\partial\boldsymbol{\mu}} = 2\mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{A} \tag{D.56}$$

by the method of undetermined coefficients.[41]

Note that equation (D.56) holds for general symmetric matrix of variables $\boldsymbol{\mu}$. In our special case in which $\boldsymbol{\mu}$ is diagonal, it is clear that $\frac{\partial\mathbb{V}(\mathrm{mc}_{it})}{\partial\boldsymbol{\mu}}$ is also diagonal, and

$$\frac{\partial\mathbb{V}(\mathrm{mc}_{it})}{\partial\boldsymbol{\mu}} = 2\left[\mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{A}\right]_{d+}, \tag{D.57}$$

where we define a linear annihilation operator $[\,\cdot\,]_{d+} : \mathbb{R}^{N\times N} \longmapsto \mathbb{R}^{N\times N}$ that restricts all off-diagonal elements to 0 while keeps the diagonal elements unchanged. Consequently,

$$\left(\frac{\partial\mathbb{V}(\mathrm{mc}_{it})}{\partial\boldsymbol{\mu}}\right)_{nn} = 2\left(\mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{A}\right)_{nn}; \qquad n = 1, 2, \ldots N$$

holds for every diagonal element.

Now, we are ready to conclude Step 1 by summarizing the properties of the Fréchet derivative for the general equilibrium system.

[41] This is indeed the First Identification Theorem that links matrix differentials to Jacobian matrices of derivatives (Magnus and Neudecker (2007); Theorem 5.6, 5.11 and Chapter 9). We also verify the correctness of formula (D.56) on the matrix calculus website: www.matrixcalculus.org. In the univariate case $N = 1$ and $\kappa = 0$, (D.56) reduces to the simple derivative $\left(\frac{\partial\mathbb{V}(\mathrm{mc}_{it})}{\partial\mu}\right) = \frac{2a\sigma^2}{(1-a\mu)^3}$.

Proposition D.2. *The Fréchet derivative for the general equilibrium system (D.39) and (D.40) in terms of attention vector $\boldsymbol{\mu}^v = diag(\boldsymbol{\mu}) \in [0,1]^N$ exists and is given by*

$$\mathcal{T}'(\boldsymbol{\mu}^v) \equiv \frac{\partial \mathcal{T}}{\partial \boldsymbol{\mu}^v} = \mathcal{T}_{\boldsymbol{\mu}^v} = \left[\text{diag}\left(\frac{\partial \mathcal{T}_1(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right) \quad \text{diag}\left(\frac{\partial \mathcal{T}_2(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right) \quad \cdots \quad \text{diag}\left(\frac{\partial \mathcal{T}_N(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right) \right]'.$$

For each sector $i = 1, 2, ...N$,

$$\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}} = \frac{\chi_i}{\theta_i \lambda_i \mathbb{V}(mc_{it})^2} \frac{\partial \mathbb{V}(mc_{it})}{\partial \boldsymbol{\mu}}; \qquad \frac{\partial \mathbb{V}(mc_{it})}{\partial \boldsymbol{\mu}} = 2\left[\mathbf{\Delta}_\mu \mathbf{\Gamma} \mathbf{\Sigma}_z \mathbf{\Gamma}' \mathbf{\Delta}_\mu' \mathbf{e}_i' \mathbf{e}_i \mathbf{\Delta}_\mu \mathbf{A}\right]_{d+},$$

where $\mathbf{\Delta}_\mu = (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}$ is the attention-distorted Leontief matrix. $[\,\cdot\,]_{d+} : \mathbb{R}^{N\times N} \longmapsto \mathbb{R}^{N\times N}$ is a linear annihilation operator that restricts all off-diagonal elements to 0 while keep the diagonal elements unchanged.[42]

Step 2: Establishing Eigenvalue Bound. By the property of spectral radius and the Gershgorin Circle Theorem,

$$\rho\left(\mathcal{T}_{\boldsymbol{\mu}^v}\right) \le \left\|\mathcal{T}_{\boldsymbol{\mu}^v}\right\|_\infty = \max_i \sum_{j=1}^N \left|\left(\mathcal{T}_{\boldsymbol{\mu}^v}\right)_{ij}\right|,$$

even if $\mathcal{T}_{\boldsymbol{\mu}^v}$ is not an non-negative matrix. By Definition 3, each row of the Jocabian matrix is simply the diagonals of $\left(\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right)$; that is,

$$\sum_{j=1}^N \left|\left(\mathcal{T}_{\boldsymbol{\mu}^v}\right)_{ij}\right| = \text{tr}\left(\left|\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right|\right); \qquad i = 1, 2, ...N,$$

where the operator $|\cdot| : \mathbb{R}^{N\times N} \longmapsto \mathbb{R}^{N\times N}$ transforms all elements of the matrix to their absolute values. Next, we show that for each sector $i = 1, 2, ...N$, the trace $\text{tr}\left(\left|\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right|\right)$ is bounded above by 1, under our sufficiency condition in Proposition 3.5.

We start with a lemma for the trace bound.

Lemma D.8. *For each sector $i = 1, 2, ...N$, the trace for the matrix derivatives after taking absolute values, $\text{tr}\left(\left|\frac{\partial \mathbb{V}(mc_{it})}{\partial \boldsymbol{\mu}}\right|\right)$, is bounded above by*

$$\text{tr}\left(\left|\frac{\partial \mathbb{V}(mc_{it})}{\partial \boldsymbol{\mu}}\right|\right) \le 2\varsigma_i; \qquad i = 1, 2...N, \tag{D.58}$$

where $\varsigma_i = \text{tr}\left(\mathbf{L}\left|\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\right|\mathbf{L}'\mathbf{e}_i'\mathbf{e}_i\mathbf{L}\mathbf{A}\right)$, $i = 1, 2, ...N$, and $\mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1}$.

Proof. For any $N \times N$ matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, it is straightforward to verify that $\boldsymbol{XY} \le |\boldsymbol{XY}| \le |\boldsymbol{X}| \cdot |\boldsymbol{Y}|$. Combined with the definition of Leontief inverse $\mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1} \ge \mathbf{\Delta}_\mu$, we derive

$$\mathbf{L}\left|\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\right|\mathbf{L}'\mathbf{e}_i'\mathbf{e}_i\mathbf{L}\mathbf{A} \ge \mathbf{\Delta}_\mu\left|\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\right|\mathbf{\Delta}_\mu'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{A} \ge \left|\mathbf{\Delta}_\mu\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_\mu'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_\mu\mathbf{A}\right| \ge 0, \tag{D.59}$$

[42] If $\boldsymbol{\kappa} = 0$, we show in the Corrolary F.1 that $\mathcal{T}_{\boldsymbol{\mu}^v} \ge 0$ is nonnegative.

since $\mathbf{L}$, $\mathbf{\Delta}_{\mu}$, $\mathbf{\Sigma}_z$, $\mathbf{e}_i$ and $\mathbf{A}$ are all nonnegative. Then, by equations (D.57) and (D.59),

$$\begin{aligned}
\operatorname{tr}\left(\left|\frac{\partial \mathbb{V}(\mathrm{mc}_{it})}{\partial \boldsymbol{\mu}}\right|\right) &= \operatorname{tr}\left(2\left|\left[\mathbf{\Delta}_{\mu}\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_{\mu}'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_{\mu}\mathbf{A}\right]_{d+}\right|\right) \\
&= 2\operatorname{tr}\left(\left|\mathbf{\Delta}_{\mu}\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\mathbf{\Delta}_{\mu}'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_{\mu}\mathbf{A}\right|\right) \\
&\leq 2\operatorname{tr}\left(\mathbf{\Delta}_{\mu}\left|\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\right|\mathbf{\Delta}_{\mu}'\mathbf{e}_i'\mathbf{e}_i\mathbf{\Delta}_{\mu}\mathbf{A}\right) \\
&\leq 2\operatorname{tr}\left(\mathbf{L}\left|\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\right|\mathbf{L}'\mathbf{e}_i'\mathbf{e}_i\mathbf{L}\mathbf{A}\right).
\end{aligned}$$

To ease notation, we define a column vector $\varsigma = \left(\{\varsigma_i\}_{i=1}^{N}\right)$, where $\varsigma_i = \operatorname{tr}\left(\mathbf{L}\left|\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\right|\mathbf{L}'\mathbf{e}_i'\mathbf{e}_i\mathbf{L}\mathbf{A}\right)$. Then the trace bound for the absolute value of matrix derivatives follows

$$\operatorname{tr}\left(\left|\frac{\partial \mathbb{V}(\mathrm{mc}_{it})}{\partial \boldsymbol{\mu}}\right|\right) \leq 2\varsigma_i; \qquad i = 1, 2 \ldots N$$

as desired. ∎

Next, we deliver an upper bound for the eigenvalues of Jacobian matrix $\mathcal{T}_{\boldsymbol{\mu}^v}$.

Proposition D.3. *Suppose for each sector $i = 1, 2, \ldots N$, we have following parameter restrictions,*

$$\kappa_i < 1, \quad \text{and} \quad \chi_i < \theta_i\lambda_i\frac{\sigma_i^4\left(1 - \varkappa_i\kappa_i\right)^4}{2\varsigma_i}, \tag{D.60}$$

where $\varkappa_i$ is defined in Proposition D.1 and $\varsigma_i = \operatorname{tr}\left(\mathbf{L}\left|\mathbf{\Gamma}\mathbf{\Sigma}_z\mathbf{\Gamma}'\right|\mathbf{L}'\mathbf{e}_i'\mathbf{e}_i\mathbf{L}\mathbf{A}\right)$ for each sector. Then the eigenvalues of the Jacobian matrix is bounded above by 1; that is

$$\rho\left(\mathcal{T}_{\boldsymbol{\mu}^v}\right) < 1. \tag{D.61}$$

Proof. Recall that each row of the Jacobian matrix is simply the diagonals of matrix of derivatives $\left(\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right)$; that is,

$$\sum_{j=1}^{N}\left|\left(\mathcal{T}_{\boldsymbol{\mu}^v}\right)_{ij}\right| = \operatorname{tr}\left(\left|\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right|\right); \qquad i = 1, 2, \ldots N.$$

Then, from equation (D.50),

$$\begin{aligned}
\operatorname{tr}\left(\left|\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right|\right) &= \frac{\chi_i}{\theta_i\lambda_i\mathbb{V}(\mathrm{mc}_{it})^2}\operatorname{tr}\left(\left|\frac{\partial \mathbb{V}(\mathrm{mc}_{it})}{\partial \boldsymbol{\mu}}\right|\right) \\
&\leq \frac{\chi_i}{\theta_i\lambda_i\left[\sigma_i^2\left(1 - \varkappa_i k_i\right)^2\right]^2}\operatorname{tr}\left(\left|\frac{\partial \mathbb{V}(\mathrm{mc}_{it})}{\partial \boldsymbol{\mu}}\right|\right) \\
&\leq \frac{\chi_i}{\theta_i\lambda_i\sigma_i^4\left(1 - \varkappa_i k_i\right)^4}\cdot 2\varsigma_i
\end{aligned}$$

$$< 1$$

holds for each sector $i = 1, 2....N$. The first inequality follows from (D.46) in Proposition D.1, the second inequality follows from (D.58), and the last inequality follows from our assumption (D.60).

Therefore, each row sum of $\mathcal{T}_{\boldsymbol{\mu}^v}$ is strictly smaller than 1, implying that

$$\left\|\mathcal{T}_{\boldsymbol{\mu}^v}\right\|_\infty = \max_i \sum_{j=1}^{N} \left|\left(\mathcal{T}_{\boldsymbol{\mu}^v}\right)_{ij}\right| < 1$$

with finitely many sectors. Now by the Gishgorin Circle Theorem

$$\rho\left(\mathcal{T}_{\boldsymbol{\mu}^v}\right) \leq \left\|\mathcal{T}_{\boldsymbol{\mu}^v}\right\|_\infty < 1$$

as desired.[43] □

Now we are ready to proceed to the final step of the proof for Proposition 3.5.

Step 3: Verify Conditions in Kellogg's Fixed Point Theorem. Let $X = \mathbb{R}^N$, which is a finite-dimensional Banach space. Let $D = (0, 1)^N$, it is clear that D is bounded open convex open subset of X. Let $\bar{D} = [0, 1]^N$ be its closure (which is also compact) and $x = \boldsymbol{\mu}^v \in \bar{D}$. The function $F(x) = \mathcal{T}(\boldsymbol{\mu}^v)$ is defined in equation (D.40). By Proposition D.3, the Fréchet derivative $F'(x) = \mathcal{T}'(\boldsymbol{\mu}^v)$ exists and is continuous. For any $\boldsymbol{\mu}^v \in D = (0, 1)^N$, Proposition D.3 states that

$$\rho\left(F'(x)\right) = \rho\left(\mathcal{T}_{\boldsymbol{\mu}^v}\right) < 1 \tag{D.62}$$

if for each sector $i = 1, 2, ...N$,

$$\kappa_i < 1, \quad \text{and} \quad \chi_i < \theta_i \lambda_i \frac{\sigma_i^4 \left(1 - \varkappa_i \kappa_i\right)^4}{2\varsigma_i}.$$

Hence 1 cannot be an eigenvalue under the assumption of Proposition 3.5. Finally, by condition (D.46) and (D.47), if

$$0 < \frac{\chi_i}{\theta_i \lambda_i \mathbb{V}(\mathrm{mc}_{it})} \leq \frac{\chi_i}{\theta_i \lambda_i \sigma_i^2 \left(1 - \varkappa_i \kappa_i\right)^2} < 1,$$

then

$$\mathcal{T}_i\left(\boldsymbol{\mu}^v\right) \in (0, 1), \forall i = 1, 2, ...N;$$

for all sector $i = 1, 2, ...N$ and all vectors $\boldsymbol{\mu}^v \in [0, 1]^N$. Therefore, for each $\boldsymbol{\mu}^v \in \partial D$ on boundary,

$$\boldsymbol{\mu}^v \neq \mathcal{T}\left(\boldsymbol{\mu}^v\right)$$

[43] If $\boldsymbol{\kappa} = 0$, then $\mathcal{T}_{\boldsymbol{\mu}^v} \geq 0$ and the Perron-Frobenious theorem leads to the same result.

Now all conditions of Kellogg's Fixed Point Theorem have been verified. It is clear that for all sectors $i = 1, 2, ...N$, if

1. The monetary policy (wage) rule satisfies $\kappa_i < 1$;
2. The information cost χ_i satisfies $0 < \chi_i < \theta_i \lambda_i \min\left\{\sigma_i^2 (1 - \varkappa_i \kappa_i)^2, \ \frac{\sigma_i^4 (1-\varkappa_i\kappa_i)^4}{2\varsigma_i}\right\}$, where $\mathbf{\Gamma} = -\mathbf{I} + \boldsymbol{\alpha\kappa}$ and $\mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1}$, the auxiliary parameters $\{\varkappa_i\}$ and $\{\varsigma_i\}$ are defined as

$$\varkappa_i = \begin{cases} 1, & 0 \le \kappa_i < 1 \\ \alpha_i, & \kappa_i < 0 \end{cases}, \quad \text{and} \quad \varsigma_i = \text{tr}\left(\mathbf{L}\left|\mathbf{\Gamma\Sigma}_z\mathbf{\Gamma}'\right|\mathbf{L}'\mathbf{e}_i'\mathbf{e}_i\mathbf{LA}\right),$$

then the general equilibrium system (3.6) has an unique fixed-point $\boldsymbol{\mu}^{v*} \in (0, 1)^N$ in the interior such that

$$\boldsymbol{\mu}^{v*} = \mathcal{T}\left(\boldsymbol{\mu}^{v*}\right)$$

The entire proof of equilibrium existence and uniqueness is now complete. □

Remark: Proposition 3.5 provides a sufficient condition that ensures equilibrium uniqueness. The condition can be further refined and relaxed in the following corrolary, utilizing the analytical form of $\mathcal{T}_{\boldsymbol{\mu}^v}$ obtained in Proposition D.2.

Corollary D.1. *Let ρ_{max} be the maximal spectral radius of $\mathcal{T}_{\boldsymbol{\mu}^v}$ in Proposition D.2 in the domain of $\boldsymbol{\mu}^v$. That is, consider an optimization problem:*

$$\rho_{max} = \max_{\boldsymbol{\mu}^v} \rho\left(\mathcal{T}_{\boldsymbol{\mu}^v}\right)$$

subject to $\mu_i \in [0, 1], \forall i = 1, 2, ...N$. If $\rho_{max} < 1$ and

$$\kappa_i < 1; \qquad 0 < \frac{\chi_i}{\theta_i \lambda_i} < \sigma_i^2 \left(1 - \varphi_i \kappa_i\right)^2; \qquad \forall i = 1, 2, ...N$$

where

$$\varsigma_i = \begin{cases} 1, & 0 \le \kappa_i < 1 \\ \alpha_i, & \kappa_i < 0 \end{cases}; \qquad i = 1, 2, ...N$$

then the general equilibrium system (3.6) has an unique fixed-point $\boldsymbol{\mu}^{v*} \in (0, 1)^N$ in the interior.

Proof. Since $\rho\left(\mathcal{T}_{\boldsymbol{\mu}^v}\right) \le \rho_{max} < 1$, (D.62) holds and 1 cannot be an eigenvalue. By condition (D.46) and (D.47), if

$$\kappa_i < 1; \qquad 0 < \frac{\chi_i}{\theta_i \lambda_i} < \sigma_i^2 (1 - \varsigma_i \kappa_i)^2; \qquad \forall i = 1, 2, ...N$$

which implies that $\mathcal{T}_i(\boldsymbol{\mu}^v) \in (0, 1)$, for all sector $i = 1, 2, ...N$ and all vectors $\boldsymbol{\mu}^v \in [0, 1]^N$. Therefore, for each $\boldsymbol{\mu}^v \in \partial D$ on boundary,

$$\boldsymbol{\mu}^v \neq \mathcal{T}\left(\boldsymbol{\mu}^v\right)$$

It is clear that all conditions of Kellogg's Fixed Point Theorem are satisfied, the general equilibrium system (3.6) has an unique fixed-point $\boldsymbol{\mu}^{v*} \in (0,1)^N$. $\square$

Remark: The optimization problem of spectral radius and attention response $\mathcal{T}$ in Corollary D.1 are independent of the equilibrium fixed-point system and can be numerically verified by Matlab. In our quantification, we verify that the calibrated parameters satisfy Corollary D.1. Therefore, the calibrated model and the model under optimal monetary policy both possess unique fixed-point in the interior.

D.9 Proof of Proposition 3.6

Proof. Recall that the households' intratemporal Euler condition, $\frac{W_t}{P_t} C_t^{-\gamma} = L_t^{\frac{1}{\eta}}$, admits the log-linearized form as

$$\ell_t = \eta(w_t - p_t^f) - \eta\gamma c_t \tag{D.63}$$

Using the derivation on nominal GDP in (D.37), consumers' budget constraint can be written as

$$c_t = w_t - p_t^f + \ell_t - \sum_{i=1}^N \lambda_i \varepsilon_{it} = \eta(w_t - p_t^f) - \gamma\eta c_t + \sum_{i=1}^N \lambda_i z_{it} = \frac{\eta}{1+\gamma\eta}(w_t - p_t^f) + \frac{1}{1+\gamma\eta}\sum_{i=1}^N \lambda_i z_{it}. \tag{D.64}$$

where the second identity follows from (D.63) and (D.38), and the last identity from collapsing the same terms of c_t. Recall that $p_t^f = \boldsymbol{\beta}'\boldsymbol{p}_t$ from the log-linearized form of the final price, then the log-linearized money supply constraint can be written as

$$m_t = p_t^f + c_t = \frac{\eta}{1+\gamma\eta} w_t + \left(1 - \frac{\eta}{1+\gamma\eta}\right) p_t^f + \frac{1}{1+\gamma\eta}\sum_{i=1}^N \lambda_i z_{it} = \frac{\eta}{1+\gamma\eta} w_t + \left(1 - \frac{\eta}{1+\gamma\eta}\right)\sum_i \beta_i p_{i,t} + \frac{1}{1+\gamma\eta}\sum_{i=1}^N \lambda_i z_{i,t}.$$

as desired. $\square$

D.10 Proof of Proposition 3.7

Proof. In the proof of Proposition F.1, we obtain the optimal posteior covariance matrix (MSE) matrix,

$$\boldsymbol{\Sigma}_{z|x_i} = \boldsymbol{\Sigma}_z - \frac{\chi_i}{2R_i d_1^i \nu_i^2}\left(\boldsymbol{\Sigma}_z \boldsymbol{\Omega}_{iz} \boldsymbol{\Sigma}_z\right),$$

and by definition $\widehat{\sigma}^2_{j|i} = \mathbb{E}\left[\left(\mathbb{E}\left[z_{jt} \mid x_{it}\right] - z_{jt}\right)^2 \mid x_{it}\right]$ is the diagonal element of the posterior matrix. If all sectors have positive level of attention such that $\mu_i > 0, \forall i = 1, 2, ...N$, we have following relations,

$$\mathrm{mc}_{it} = R_i^{-\frac{1}{2}} \mathbf{G}_i \boldsymbol{z}_t = \frac{1}{\mu_i} p_{it}$$

Therefore, when $\Sigma_{\mathbf{z}}$ is diagonal, the diagonal element of posterior is given by

$$\widehat{\sigma}^2_{j|i} = \sigma_j^2 - \frac{\chi_i}{2d_1^i \nu_i^2} \frac{1}{\mu_i^2} \phi_{ij}^2 \sigma_j^4$$

where we use the definition of weighting matrix $\mathbf{\Omega}_{iz}$. Applying our previous results, it is easy to deduce that $\frac{\chi_i}{2d_1^i \nu_i^2} = \frac{1}{\mathbb{V}(mc_{it})+\nu_i^2}$. It then follows that

$$\widehat{\sigma}^2_{j|i} = \sigma_j^2 - \frac{\phi_{ij}^2 \sigma_j^4}{\mu_i^2 \left(\mathbb{V}\left(mc_{it}\right) + \nu_i^2\right)} = \sigma_j^2 - \frac{\mu_i \phi_{ij}^2 \sigma_j^4}{\sum_{k=1}^N \sigma_k^2 \phi_{ik}^2}$$

where I use the fact that $\mu_i = \frac{\mathbb{V}(mc_{it})}{\mathbb{V}(mc_{it})+\nu_i^2}$ and $\mathbb{V}\left(p_{it}\right) = \mu_i^2 \mathbb{V}\left(mc_{it}\right)$. Finally, the relative attention allocation measure is given by

$$\omega_{ij} = \frac{\sigma_j^2 - \widehat{\sigma}^2_{j|i}}{\sigma_j^2} = \mu_i \frac{\phi_{ij}^2 \sigma_j^2}{\sum_{k=1}^N \sigma_k^2 \phi_{ik}^2}$$

By construction, the fraction $\frac{\phi_{ij}^2 \sigma_j^2}{\sum_{k=1}^N \sigma_k^2 \phi_{ik}^2}$ is a well-defined probability measure,

$$\frac{\phi_{ij}^2 \sigma_j^2}{\sum_{k=1}^N \sigma_k^2 \phi_{ik}^2} \in [0,1]; \qquad \sum_{j=1}^N \frac{\phi_{ij}^2 \sigma_j^2}{\sum_{k=1}^N \sigma_k^2 \phi_{ik}^2} = 1; \qquad \forall i = 1, 2, ..N.$$

Therefore, we have $\sum_{j=1}^N \omega_{ij} = \mu_i$ as desired. The proof is complete. □

E. Optimal Monetary Policy

E.1 Proof of Lemma 4.1

Proof. To study the optimal policy problem, we define such loss as difference of the representative household's utilities in the RI economy and in a reference economy with perfect information. In what follows, variables in the full-information (FI) economy are denoted with *.

Following La'O and Tahbaz-Salehi (2022), we first derive the second-order approximation of the welfare function as

$$\mathbb{E}\left[U_t - U_t^*\right] \approx -\frac{1}{2}\left[\sum_{i=1}^{N} \lambda_i \theta_i \mathcal{D}_i + (\gamma + 1/\eta)\mathbb{V}\left(c_t - c_t^*\right) + \sum_{i=0}^{N} \lambda_i C_i\right], \tag{E.1}$$

where $\{C_i, \mathcal{D}_i\}_{i=1}^N$ are defined in the Lemma. The output-gap volatility is defined as,

$$\mathbb{V}\left(c_t - c_t^*\right) = \frac{1}{(\gamma + 1/\eta)^2}\mathbb{E}\left(\sum_{j=1}^{n} \beta_j \bar{e}_{jt}\right)^2 \tag{E.2}$$

U_t and U_t^* are the utilities in the RI and FI economies, respectively.[44] The expectation operator in (4.1) is defined with respect to the central bank (CB)'s information set whose information is assumed to be perfect. For brevity, we omit the algebraic details of this second-order approximation.

Using the equilibrium optimal signal structure we derive in Section 3, the within-sector pricing error variation in sector i is given by

$$\mathcal{D}_i = \int_0^1 \left(p_{\iota t} - p_{it}\right)^2 d\iota = \mu_i^2 \mathbb{E} u_{\iota t}^2 = \mu_i^2 \nu_i^2 = \frac{\chi_i}{\theta_i \lambda_i}\mu_i \tag{E.3}$$

where the second and third equality follows from the fact that firms within sector i solve a identical information acquisition problem with endogenous noises drawn from the same distribution: $u_{\iota t}^2 \sim N(0, \nu_i^2)$. The last equality follows from the optimal solution for noise variance,

$$\nu_i^2 = \frac{\frac{\chi_i}{\theta_i \lambda_i}\mathbb{V}(\mathrm{mc}_{it})}{\mathbb{V}(\mathrm{mc}_{it}) - \frac{\chi_i}{\theta_i \lambda_i}} = \frac{1}{\mu_i}\frac{\chi_i}{\theta_i \lambda_i}; \tag{E.4}$$

(E.3) implies the variations of within-sector pricing errors are determined by the nominal rigidity μ_i and the variance of endogenous noise ν_i^2. The term $\mathcal{D}_i$ vanishes when information is perfect ($\nu_i^2 = 0$), or when information is completely rigid ($\mu_i = 0$). If firms collect information subject to capacity constraint, price flexibility and noise variance move in the opposite direction – higher attentiveness is accompanied by lower variances, with their product equals to a constant $\frac{\chi_i}{\theta_i \lambda_i} > 0$. If the information cost $\chi_i > 0$ is strictly positive, the minimally attainable

[44] Both economies are approximated around the same deterministic steady-state. In the flexible-price economy, price levels are indeterminate. We normalize wage such that $W_t^* = W_t$, where W_t in the RI economy is uniquely pinned down by the monetary policy.

noise is achieved at the maximum level of price flexibility $\mu_i = 1$ as the volatility of profit-maximizing price converges to infinity,

$$\lim_{\mathbb{V}(\mathrm{mc}_{it})\to\infty} \nu_i^2 = \lim_{\mu_i \to 1} \nu_i^2 = \frac{\chi_i}{\theta_i \lambda_i} > 0 \tag{E.5}$$

That is, even with fully-flexible price, full-information equilibrium cannot be restored given the positive level of noise in RI equilibrium. Meanwhile, (E.3) and (E.4) indicate the price flexibility μ_i plays a dominant role in driving within-sector pricing-error dispersion. □

E.2 Proof of Lemma 4.2

Proof. Consider the objective function in (4.1), we express the objective in matrix form,

$$\mathbb{E}\left[U_t - U_t^*\right] = -\frac{1}{2}\left[\boldsymbol{\chi}^T \boldsymbol{\mu}^v + \frac{1}{(\gamma + 1/\eta)}\boldsymbol{\beta}^T \boldsymbol{\Sigma}_e \boldsymbol{\beta} + \boldsymbol{\lambda}^T \operatorname{diag}\left(\boldsymbol{\Sigma}_e\right) - \boldsymbol{\lambda}^T \operatorname{diag}\left(A\boldsymbol{\Sigma}_e A^T\right) - \boldsymbol{\beta}^T \boldsymbol{\Sigma}_e \boldsymbol{\beta}\right] \tag{E.6}$$

where we use the expression in (3.9) to define the covariance matrix of the pricing errors as

$$\boldsymbol{\Sigma}_e = \mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{\Sigma}_z\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right)' \mathbf{Q}' \geq 0.$$

The central bank's goal is to maximize the expected welfare under rational inattention, or equivalently, minimize the expected welfare loss due to RI. Therefore, we define the CB's optimal policy problem formally as

Definition 4. *The central bank designs optimal monetary policy by solving the following constrained optimization problem*

$$\min_{\{\boldsymbol{\mu}^v, \boldsymbol{\kappa}\}} \Delta U_t^{within} + \Delta U_t^{OG} + \Delta U_t^{across} = \frac{1}{2}\left[\boldsymbol{\chi}^T \boldsymbol{\mu}^v + \frac{1}{(\gamma + 1/\eta)}\boldsymbol{\beta}^T \boldsymbol{\Sigma}_e \boldsymbol{\beta} + \boldsymbol{\lambda}^T \operatorname{diag}\left(\boldsymbol{\Sigma}_e\right) - \boldsymbol{\lambda}^T \operatorname{diag}\left(\mathbf{A}\boldsymbol{\Sigma}_e \mathbf{A}^T\right) - \boldsymbol{\beta}^T \boldsymbol{\Sigma}_e \boldsymbol{\beta}\right]$$

subject to the equilibrium fixed-point for each sector $i = 1, 2 \ldots N$,

$$\mu_i = \max\left\{0, 1 - \frac{\chi_i}{\theta_i \lambda_i \mathbb{V}(mc_{it})}\right\}; \quad \mathbb{V}(mc_{it}) = \left\|\mathbf{e}_i\left[\left(\mathbf{I} - \mathbf{A}\boldsymbol{\mu}\right)^{-1}\left(-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa}\right)\right]\boldsymbol{\Sigma}_z^{1/2}\right\|^2. \tag{E.7}$$

The proof is now complete. □

E.3 Proof of Proposition 4.1

Proof. By Proposition 3.3, the optimal signal structure under fixed-capacity is identical to the elastic attention model. The sectoral nominal rigidities/attentions are exogenously given by $\mu_i = 1 - e^{-2\delta_i} = \frac{1}{1+\tau_i}$, $\forall i = 1, 2, \ldots N$. Under welfare objective function (E.6), we only need to update the expression for the first term of welfare loss. By (E.3), the within-sector cross-section dispersion of pricing errors are given by

$$\mathcal{D}_i = \mathbb{E}\left[\int_0^1 \left(p_{\iota i t} - p_{it}\right)^2 d\iota\right] = \mu_i^2 \mathbb{E} u_{\iota i t}^2 = \mu_i^2 \nu_i^2$$

Using the results in the proof of Proposition 3.3, we note that an important feature of fixed-capacity model is that the variance of endogenous noise is proportional to the volatility of marginal cost:

$$\nu_i^2 = R_i^{-1} \frac{e^{-2\kappa_i}}{1 - e^{-2\kappa_i}} \left\| \Sigma_{\mathbf{z}}^{\frac{1}{2}} \mathbf{G}_i' \right\|^2 = \frac{e^{-2\kappa_i}}{1 - e^{-2\kappa_i}} \mathbb{V}\left(\mathrm{mc}_{it}\right). \tag{E.8}$$

It then follows from the above equation

$$\mathcal{D}_i = \mu_i (1 - \mu_i) \mathbb{V}\left(\mathrm{mc}_{it}\right). \tag{E.9}$$

On the other hand, it follows from (D.31) and (D.32) that under exogenous information, the within-sector cross-section dispersion of pricing errors are given by

$$\mathcal{D}_i = \frac{1}{(1+\tau_i)^2} \int \sum_{j=1}^{N} m_j^2 u_{\iota ijt}^2 d\iota = \frac{\tau_i}{(1+\tau_i)^2} \sum_{j=1}^{N} m_j^2 \sigma_j^2 = \mu_i (1 - \mu_i) \mathbb{V}\left(\mathrm{mc}_{it}\right) \tag{E.10}$$

Therefore, (E.10) and (E.9) coincides. More generally, given exogenous $\boldsymbol{\mu}$, the welfare loss functions under two cases are identical, given by (4.1). Since the attention channel of monetary policy is muted in these two cases, the unconstrained optimal monetary policy coincides in these two economies. The solution to the optimal policy is straightforward. The only updated component is given by the derivative of the within-sector welfare loss with respect to sectoral wage rule κ_s,

$$\frac{\partial \Delta U_t^{within}}{\partial \kappa_s} = \frac{d\mathbb{E}\left[\sum_{i=1}^{N} \lambda_i \theta_i \vartheta_{it}\right]}{d\kappa_s} = \sum_{i=1}^{N} \lambda_i \theta_i \mu_i \left(1 - \mu_i\right) \frac{dV_i}{d\kappa_s} \tag{E.11}$$

where we define $V_i \equiv \mathbb{V}\left(\mathrm{mc}_{it}\right)$ to simplify notation. In the Proof of Proposition 4.2, (E.37), we express the following matrix derivative as

$$\begin{bmatrix} \frac{\partial V_1(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \\ \frac{\partial V_2(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \\ \vdots \\ \frac{\partial V_N(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \end{bmatrix} = 2\,\mathrm{diag}\left(\boldsymbol{\rho}\right) \boldsymbol{\mu}^{-1} \left(\mathbf{I} - \mathbf{Q}\right) \left(-\mathbf{L} + \mathbf{1}\boldsymbol{\kappa}\right) \boldsymbol{\Sigma}_{\mathbf{z}} \tag{E.12}$$

Using the techniques developed in the proof of Lemma 4.3, Proposition E.2, and Proposition 4.3, we write the first order condition w.r.t $\boldsymbol{\kappa}$ in row:

$$\left[\boldsymbol{\lambda}'\,\mathrm{diag}\left(\left\{\theta_i \rho_i \left(1 - \mu_i\right)\right\}_{i=1}^{N}\right) \left(\mathbf{I} - \mathbf{Q}\right) + \left\{ \left(\frac{1}{(\gamma + 1/\eta)} - 1\right) \left(\boldsymbol{\beta}'\mathbf{Q}\mathbf{1}\right) \boldsymbol{\beta}' + \boldsymbol{\lambda}' \Big(\mathrm{diag}\left(\mathbf{Q}\mathbf{1}\right) - \mathrm{diag}\left(\mathbf{A}\mathbf{Q}\mathbf{1}\right) \mathbf{A} \Big) \right\} \mathbf{Q} \right] \left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right) = \mathbf{0}'$$

Then we obtain the optimal policy weight as

$$\varphi = \boldsymbol{\lambda}' \operatorname{diag}\left(\left\{\theta_i \rho_i \left(1-\mu_i\right)\right\}_{i=1}^{N}\right) + \left\{\left(\frac{1}{(\gamma+1/\eta)}-1\right)(\boldsymbol{\beta}'\mathbf{Q}\mathbf{1})\,\boldsymbol{\beta}' + \boldsymbol{\lambda}'\Big(\operatorname{diag}(\mathbf{Q}\mathbf{1}) - \operatorname{diag}(\mathbf{A}\mathbf{Q}\mathbf{1})\,\mathbf{A}\Big)\right\}\mathbf{L}\left(\boldsymbol{\mu}^{-1}-\mathbf{I}\right) \tag{E.13}$$

where we use the fact that $\boldsymbol{\phi} = (\mathbf{I}-\mathbf{Q})(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})$. Written in scalars,

$$\varphi_i^{exo} = \varphi_i^{fix} = \left[\mu_i\lambda_i\theta_i\rho_i + \frac{(1-\rho_0)}{(\gamma+1/\eta)}\lambda_i + \sum_{j=1}^{N}\left(1-\mu_i\right)\lambda_j\rho_j l_{ji} + \left(\rho_0-\rho_i\right)\lambda_i\right]\left(\frac{1}{\mu_i}-1\right) \tag{E.14}$$

for each sector $i = 1, 2, ..N$. We refer readers to the proof of Lemma 4.3, Proposition E.2, and Proposition 4.3 for the transformation of the price-stabilization and the algebra details that lead to (E.14). The proof is now complete. □

E.4 Proof of Proposition 4.2

Proof. In the fixed-point system (3.6), the implicit functions of $\boldsymbol{\mu}$ and $\boldsymbol{\kappa}$ can be expressed as

$$\mu_i = \mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa}); \qquad i = 1, 2, ...N, \tag{E.15}$$

where $\mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa})$ is a scalar best response function of $\boldsymbol{\mu}$ and $\boldsymbol{\kappa}$,

$$\mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa}) = 1 - \frac{\chi_i}{\theta_i\lambda_i V_i(\boldsymbol{\mu}, \boldsymbol{\kappa})}; \qquad V_i(\boldsymbol{\mu}, \boldsymbol{\kappa}) = \mathbb{V}(\mathrm{mc}_{it}); \qquad i = 1, 2, ...N. \tag{E.16}$$

Take differential on both sides of (E.15),

$$d\mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa}) = d\mu_i; \qquad i = 1, 2, ...N.$$

We write $\mathcal{T}_i(\boldsymbol{\mu}, \boldsymbol{\kappa})$ briefly as $\mathcal{T}_i$, then $\frac{d\mathcal{T}_i}{d\boldsymbol{\kappa}} = \frac{d\mu_i}{d\boldsymbol{\kappa}}$; $i = 1, 2, ...N$.. In matrix form,

$$\begin{bmatrix} \frac{d\mathcal{T}_1}{d\boldsymbol{\kappa}'} & \frac{d\mathcal{T}_2}{d\boldsymbol{\kappa}'} & \cdots & \frac{d\mathcal{T}_N}{d\boldsymbol{\kappa}'} \end{bmatrix} = \begin{bmatrix} \frac{d\mu_1}{d\boldsymbol{\kappa}'} & \frac{d\mu_2}{d\boldsymbol{\kappa}'} & \cdots & \frac{d\mu_N}{d\boldsymbol{\kappa}'} \end{bmatrix}. \tag{E.17}$$

In the best response function (E.15), given $\boldsymbol{\kappa}$, the differential of μ_i can be written as

$$d\mu_i = \sum_{j=1}^{N}\frac{\partial\mu_i}{\partial\kappa_j}d\kappa_j; \qquad i = 1, 2, ...N,$$

which implies that

$$\frac{d\mu_i}{d\kappa_s} = \frac{\partial\mu_i}{\partial\kappa_s}; \qquad i, s = 1, 2, ...N. \tag{E.18}$$

Now consider differential of the best response function $\mathcal{T}_i$,

$$d\mathcal{T}_i = \left(\frac{\partial \mathcal{T}_i}{\partial \mu_1} d\mu_1 + \frac{\partial \mathcal{T}_i}{\partial \mu_2} d\mu_2 + \cdots \frac{\partial \mathcal{T}_i}{\partial \mu_N} d\mu_N\right) + \left(\frac{\partial \mathcal{T}_i}{\partial \kappa_1} d\kappa_1 + \frac{\partial \mathcal{T}_i}{\partial \kappa_2} d\kappa_2 + \cdots \frac{\partial \mathcal{T}_i}{\partial \kappa_N} d\kappa_N\right); \qquad i = 1, 2, ...N. \tag{E.19}$$

Dividing both sides of equation (E.19) by each sector's κ_i, respectively,

$$\begin{aligned} \frac{d\mathcal{T}_i}{d\kappa_1} &= \left(\frac{\partial \mathcal{T}_i}{\partial \mu_1}\frac{d\mu_1}{d\kappa_1} + \frac{\partial \mathcal{T}_i}{\partial \mu_2}\frac{d\mu_2}{d\kappa_1} + \cdots \frac{\partial \mathcal{T}_i}{\partial \mu_N}\frac{d\mu_N}{d\kappa_1}\right) + \left(\frac{\partial \mathcal{T}_i}{\partial \kappa_1}\right) \\ \frac{d\mathcal{T}_i}{d\kappa_2} &= \left(\frac{\partial \mathcal{T}_i}{\partial \mu_1}\frac{d\mu_1}{d\kappa_2} + \frac{\partial \mathcal{T}_i}{\partial \mu_2}\frac{d\mu_2}{d\kappa_2} + \cdots \frac{\partial \mathcal{T}_i}{\partial \mu_N}\frac{d\mu_N}{d\kappa_2}\right) + \left(\frac{\partial \mathcal{T}_i}{\partial \kappa_2}\right) \\ &\cdots \\ \frac{d\mathcal{T}_i}{d\kappa_N} &= \left(\frac{\partial \mathcal{T}_i}{\partial \mu_1}\frac{d\mu_1}{d\kappa_N} + \frac{\partial \mathcal{T}_i}{\partial \mu_2}\frac{d\mu_2}{d\kappa_N} + \cdots \frac{\partial \mathcal{T}_i}{\partial \mu_N}\frac{d\mu_N}{d\kappa_N}\right) + \left(\frac{\partial \mathcal{T}_i}{\partial \kappa_N}\right). \end{aligned} \tag{E.20}$$

Consolidating (E.20) into matrix form and combining (E.18), we get

$$\frac{d\mathcal{T}_i}{d\boldsymbol{\kappa}'} = \begin{bmatrix} \frac{\partial \mu_1}{\partial \kappa_1} & \frac{\partial \mu_2}{\partial \kappa_1} & \cdots & \frac{\partial \mu_N}{\partial \kappa_1} \\ \frac{\partial \mu_1}{\partial \kappa_2} & \frac{\partial \mu_2}{\partial \kappa_2} & \cdots & \frac{\partial \mu_N}{\partial \kappa_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mu_1}{\partial \kappa_N} & \frac{\partial \mu_2}{\partial \kappa_N} & \cdots & \frac{\partial \mu_N}{\partial \kappa_N} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{T}_i}{\partial \mu_1} \\ \frac{\partial \mathcal{T}_i}{\partial \mu_2} \\ \vdots \\ \frac{\partial \mathcal{T}_i}{\partial \mu_N} \end{bmatrix} + \begin{bmatrix} \frac{\partial \mathcal{T}_i}{\partial \kappa_1} \\ \frac{\partial \mathcal{T}_i}{\partial \kappa_2} \\ \vdots \\ \frac{\partial \mathcal{T}_i}{\partial \kappa_N} \end{bmatrix},$$

More specifically,

$$\begin{bmatrix} \frac{d\mathcal{T}_1}{d\boldsymbol{\kappa}'} & \frac{d\mathcal{T}_2}{d\boldsymbol{\kappa}'} & \cdots & \frac{d\mathcal{T}_N}{d\boldsymbol{\kappa}'} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mu_1}{\partial \kappa_1} & \frac{\partial \mu_2}{\partial \kappa_1} & \cdots & \frac{\partial \mu_N}{\partial \kappa_1} \\ \frac{\partial \mu_1}{\partial \kappa_2} & \frac{\partial \mu_2}{\partial \kappa_2} & \cdots & \frac{\partial \mu_N}{\partial \kappa_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mu_1}{\partial \kappa_N} & \frac{\partial \mu_2}{\partial \kappa_N} & \cdots & \frac{\partial \mu_N}{\partial \kappa_N} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{T}_1}{\partial \mu_1} & \frac{\partial \mathcal{T}_2}{\partial \mu_1} & \cdots & \frac{\partial \mathcal{T}_N}{\partial \mu_1} \\ \frac{\partial \mathcal{T}_1}{\partial \mu_2} & \frac{\partial \mathcal{T}_2}{\partial \mu_2} & \cdots & \frac{\partial \mathcal{T}_N}{\partial \mu_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mathcal{T}_1}{\partial \mu_N} & \frac{\partial \mathcal{T}_2}{\partial \mu_N} & \cdots & \frac{\partial \mathcal{T}_N}{\partial \mu_N} \end{bmatrix} + \begin{bmatrix} \frac{\partial \mathcal{T}_1}{\partial \kappa_1} & \frac{\partial \mathcal{T}_2}{\partial \kappa_1} & \cdots & \frac{\partial \mathcal{T}_N}{\partial \kappa_1} \\ \frac{\partial \mathcal{T}_1}{\partial \kappa_2} & \frac{\partial \mathcal{T}_2}{\partial \kappa_2} & \cdots & \frac{\partial \mathcal{T}_N}{\partial \kappa_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mathcal{T}_1}{\partial \kappa_N} & \frac{\partial \mathcal{T}_2}{\partial \kappa_N} & \cdots & \frac{\partial \mathcal{T}_N}{\partial \kappa_N} \end{bmatrix} \tag{E.21}$$

We define two canonical Jacobian matrices of $\mathcal{T}_i$ with respect to $\boldsymbol{\mu}$ and $\boldsymbol{\kappa}$, respectively as

$$\mathcal{T}_{\boldsymbol{\mu}^v} \equiv \frac{\partial \mathcal{T}}{\partial \boldsymbol{\mu}^v} = \begin{bmatrix} \frac{\partial \mathcal{T}_1}{\partial \mu_1} & \frac{\partial \mathcal{T}_1}{\partial \mu_2} & \cdots & \frac{\partial \mathcal{T}_1}{\partial \mu_N} \\ \frac{\partial \mathcal{T}_2}{\partial \mu_1} & \frac{\partial \mathcal{T}_2}{\partial \mu_2} & \cdots & \frac{\partial \mathcal{T}_2}{\partial \mu_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mathcal{T}_N}{\partial \mu_1} & \frac{\partial \mathcal{T}_N}{\partial \mu_2} & \cdots & \frac{\partial \mathcal{T}_N}{\partial \mu_N} \end{bmatrix}, \quad \text{and} \quad \mathcal{T}_{\boldsymbol{\kappa}} \equiv \frac{\partial \mathcal{T}}{\partial \boldsymbol{\kappa}} = \begin{bmatrix} \frac{\partial \mathcal{T}_1}{\partial \kappa_1} & \frac{\partial \mathcal{T}_1}{\partial \kappa_2} & \cdots & \frac{\partial \mathcal{T}_1}{\partial \kappa_N} \\ \frac{\partial \mathcal{T}_2}{\partial \kappa_1} & \frac{\partial \mathcal{T}_2}{\partial \kappa_2} & \cdots & \frac{\partial \mathcal{T}_2}{\partial \kappa_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mathcal{T}_N}{\partial \kappa_1} & \frac{\partial \mathcal{T}_N}{\partial \kappa_2} & \cdots & \frac{\partial \mathcal{T}_N}{\partial \kappa_N} \end{bmatrix}, \tag{E.22}$$

where $\mathcal{T}_{\boldsymbol{\mu}^v}$ is identical to the definition in euqation (D.48). Similarly, we define the canonical Jacobian matrix of $\boldsymbol{\mu}^v$ with respect to $\boldsymbol{\kappa}$ as,

$$J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa}) \equiv \frac{d\boldsymbol{\mu}^v}{d\boldsymbol{\kappa}} = \begin{bmatrix} \frac{\partial\mu_1}{\partial\kappa_1} & \frac{\partial\mu_1}{\partial\kappa_2} & \cdots & \frac{\partial\mu_1}{\partial\kappa_N} \\ \frac{\partial\mu_2}{\partial\kappa_1} & \frac{\partial\mu_2}{\partial\kappa_2} & \cdots & \frac{\partial\mu_2}{\partial\kappa_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial\mu_N}{\partial\kappa_1} & \frac{\partial\mu_2}{\partial\kappa_2} & \cdots & \frac{\partial\mu_N}{\partial\kappa_N} \end{bmatrix}, \tag{E.23}$$

which is defined implicitly by the equilibrium fixed-point system. Note the above equation contains all the information regarding the dependence of the diagonal matrix $\boldsymbol{\mu}$ on $\boldsymbol{\kappa}$. Using the above notations, we express equation (E.21) as

$$\begin{bmatrix} \frac{d\mathcal{T}_1}{d\boldsymbol{\kappa}'} & \frac{d\mathcal{T}_2}{d\boldsymbol{\kappa}'} & \cdots & \frac{d\mathcal{T}_N}{d\boldsymbol{\kappa}'} \end{bmatrix} = J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})' \mathcal{T}'_{\boldsymbol{\mu}^v} + \mathcal{T}'_{\boldsymbol{\kappa}}.$$

On the other hand, we rearrange equation (E.18) into matrix form,

$$\begin{bmatrix} \frac{d\mu_1}{d\boldsymbol{\kappa}'} & \frac{d\mu_2}{d\boldsymbol{\kappa}'} & \cdots & \frac{d\mu_N}{d\boldsymbol{\kappa}'} \end{bmatrix} = \begin{bmatrix} \frac{\partial\mu_1}{\partial\kappa_1} & \frac{\partial\mu_2}{\partial\kappa_1} & \cdots & \frac{\partial\mu_N}{\partial\kappa_1} \\ \frac{\partial\mu_1}{\partial\kappa_2} & \frac{\partial\mu_2}{\partial\kappa_2} & \cdots & \frac{\partial\mu_N}{\partial\kappa_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial\mu_1}{\partial\kappa_N} & \frac{\partial\mu_2}{\partial\kappa_N} & \cdots & \frac{\partial\mu_N}{\partial\kappa_N} \end{bmatrix} = J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})'. \tag{E.24}$$

Substituting (E.23) and (E.24) into (E.17) yields

$$J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})' \mathcal{T}'_{\boldsymbol{\mu}^v} + \mathcal{T}'_{\boldsymbol{\kappa}} = J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})'.$$

Therefore,

$$J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa}) \equiv \frac{d\boldsymbol{\mu}^v}{d\boldsymbol{\kappa}} = \left[\mathbf{I} - \mathcal{T}_{\boldsymbol{\mu}^v}\right]^{-1} \mathcal{T}_{\boldsymbol{\kappa}}, \tag{E.25}$$

As desired. Next, we proceed to solve the analytical form of two Jacobian matrices $\mathcal{T}_{\boldsymbol{\mu}^v}$ and $\mathcal{T}_{\boldsymbol{\kappa}}$. Define an operator $[\,\cdot\,]_{(i,j)} : \mathbb{R}^{m\times n} \longmapsto \mathbb{R}$ that extracts the ijth element of any $m \times n$ matrix—a notation similar to computer programming. Similarly, $[\,\cdot\,]_{(i,:)} : \mathbb{R}^{m\times n} \longmapsto \mathbb{R}^{1\times n}$ extracts the i row, and $[\,\cdot\,]_{(:,j)} : \mathbb{R}^{m\times n} \longmapsto \mathbb{R}^{m\times 1}$ extracts the j column of matrices. The next Lemma presents a matrix transformation result.

Lemma E.1. *Given a N-dimensional row vector $\mathbf{v}$ and two $N \times N$ matrices $\mathbf{X}$ and $\mathbf{Y}$, it follows that*

$$\begin{bmatrix} \mathbf{v}\mathbf{e}'_1\mathbf{e}_1\mathbf{X} \\ \mathbf{v}\mathbf{e}'_2\mathbf{e}_2\mathbf{X} \\ \vdots \\ \mathbf{v}\mathbf{e}'_N\mathbf{e}_N\mathbf{X} \end{bmatrix} = \operatorname{diag}(\mathbf{v}')\mathbf{X}; \quad \textit{and} \quad \begin{bmatrix} \left(\operatorname{diag}\left(\mathbf{X}\mathbf{e}'_1\mathbf{e}_1\mathbf{Y}\right)\right)' \\ \left(\operatorname{diag}\left(\mathbf{X}\mathbf{e}'_2\mathbf{e}_2\mathbf{Y}\right)\right)' \\ \vdots \\ \left(\operatorname{diag}\left(\mathbf{X}\mathbf{e}'_N\mathbf{e}_N\mathbf{Y}\right)\right)' \end{bmatrix} = \mathbf{X}' \odot \mathbf{Y},$$

where $\odot$ denotes the Hadamard product, and $\mathbf{e}_i$ denotes the ith standard basis (row) vector.

Proof. By definition,

$$\begin{bmatrix} \mathbf{v}\mathbf{e}_1'\mathbf{e}_1\mathbf{X} \\ \mathbf{v}\mathbf{e}_2'\mathbf{e}_2\mathbf{X} \\ \vdots \\ \mathbf{v}\mathbf{e}_N'\mathbf{e}_N\mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{v}\mathbf{e}_1'\mathbf{e}_1 \\ \mathbf{v}\mathbf{e}_2'\mathbf{e}_2 \\ \vdots \\ \mathbf{v}\mathbf{e}_N'\mathbf{e}_N \end{bmatrix} \mathbf{X} = \text{diag}(\mathbf{v}')\mathbf{X}.$$

Next, we observe that

$$\left(\text{diag}\left(\mathbf{X}\mathbf{e}_i'\mathbf{e}_i\mathbf{Y}\right)\right)' = \left([\mathbf{X}]_{(:,i)}\right)' \odot [\mathbf{Y}]_{(i,:)} = [\mathbf{X}']_{(i,:)} \odot [\mathbf{Y}]_{(i,:)}; \qquad \forall i = 1, 2, ...N$$

leading to

$$\begin{bmatrix} \left(\text{diag}\left(\mathbf{X}\mathbf{e}_1'\mathbf{e}_1\mathbf{Y}\right)\right)' \\ \left(\text{diag}\left(\mathbf{X}\mathbf{e}_2'\mathbf{e}_2\mathbf{Y}\right)\right)' \\ \vdots \\ \left(\text{diag}\left(\mathbf{X}\mathbf{e}_N'\mathbf{e}_N\mathbf{Y}\right)\right)' \end{bmatrix} = \begin{bmatrix} [\mathbf{X}']_{(1,:)} \odot [\mathbf{Y}]_{(1,:)} \\ [\mathbf{X}']_{(2,:)} \odot [\mathbf{Y}]_{(2,:)} \\ \vdots \\ [\mathbf{X}']_{(N,:)} \odot [\mathbf{Y}]_{(N,:)} \end{bmatrix} = \mathbf{X}' \odot \mathbf{Y}.$$

□

Now recall from Proposition D.2, the Jacobian matirx $\mathcal{T}_{\boldsymbol{\mu}^v}$ is given by

$$\mathcal{T}_{\boldsymbol{\mu}^v} = \left[\text{diag}\left(\frac{\partial \mathcal{T}_1(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right) \quad \text{diag}\left(\frac{\partial \mathcal{T}_2(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right) \quad \cdots \quad \text{diag}\left(\frac{\partial \mathcal{T}_N(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right)\right]'. \tag{E.26}$$

when we treat $\boldsymbol{\kappa}$ as a coefficient row vector. For each sector $i = 1, 2, ...N$,

$$\frac{\partial \mathcal{T}_i(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}} = \frac{\chi_i}{\theta_i \lambda_i \mathbb{V}(\text{mc}_{it})^2} \frac{\partial \mathbb{V}(\text{mc}_{it})}{\partial \boldsymbol{\mu}}; \qquad \frac{\partial \mathbb{V}(\text{mc}_{it})}{\partial \boldsymbol{\mu}} = 2\left[\boldsymbol{\Delta}_{\boldsymbol{\mu}} \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z \boldsymbol{\Gamma}' \boldsymbol{\Delta}_{\boldsymbol{\mu}}' \mathbf{e}_i' \mathbf{e}_i \boldsymbol{\Delta}_{\boldsymbol{\mu}} \mathbf{A}\right]_{d+}.$$

We transform (E.26) using Lemma E.1 as

$$\begin{aligned} \mathcal{T}_{\boldsymbol{\mu}^v} = \begin{bmatrix} \left[\text{diag}\left(\frac{\partial \mathcal{T}_1(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right)\right]' \\ \left[\text{diag}\left(\frac{\partial \mathcal{T}_2(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right)\right]' \\ \vdots \\ \left[\text{diag}\left(\frac{\partial \mathcal{T}_N(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right)\right]' \end{bmatrix} &= 2\,\text{diag}\left(\left\{\frac{\chi_i}{\theta_i \lambda_i V_i^2}\right\}_{i=1}^N\right) \begin{bmatrix} \left(\text{diag}\left(\boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z \boldsymbol{\Gamma}' \boldsymbol{\Delta}_\mu' \mathbf{e}_1' \mathbf{e}_1 \boldsymbol{\Delta}_\mu \mathbf{A}\right)\right)' \\ \left(\text{diag}\left(\boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z \boldsymbol{\Gamma}' \boldsymbol{\Delta}_\mu' \mathbf{e}_2' \mathbf{e}_2 \boldsymbol{\Delta}_\mu \mathbf{A}\right)\right)' \\ \vdots \\ \left(\text{diag}\left(\boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z \boldsymbol{\Gamma}' \boldsymbol{\Delta}_\mu' \mathbf{e}_N' \mathbf{e}_N \boldsymbol{\Delta}_\mu \mathbf{A}\right)\right)' \end{bmatrix} \\ &= 2\,\text{diag}\left(\left\{\frac{\chi_i}{\theta_i \lambda_i V_i^2}\right\}_{i=1}^N\right) \left[\left(\boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z \boldsymbol{\Gamma}' \boldsymbol{\Delta}_\mu'\right) \odot \left(\boldsymbol{\Delta}_\mu \mathbf{A}\right)\right]. \end{aligned} \tag{E.27}$$

where we simplify the notation $V_i = \mathbb{V}(\text{mc}_{it})$. By equation (D.36) and the definition of $\boldsymbol{\phi} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1} \boldsymbol{\mu}(-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa})$, we obtain a matrix identity

$$\boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} = (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}(-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa}) = \boldsymbol{\mu}^{-1}(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}(-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa}) = \boldsymbol{\mu}^{-1}\boldsymbol{\phi}. \tag{E.28}$$

Combining the above identities with the definition of covariance matrix of the vector of sectoral marginal cost,

$$\mathbb{V}\left(\mathbf{mc}_t\right) \equiv \Sigma_{\mathbf{p}_t^{\Delta}} = \boldsymbol{\mu}^{-1}\boldsymbol{\phi}\boldsymbol{\Sigma}_z\boldsymbol{\phi}'\boldsymbol{\mu}^{-1} = \mathbb{COV}\left(\mathbf{mc}_t, \mathbf{mc}_t\right) \tag{E.29}$$

and $\boldsymbol{\Delta}_{\boldsymbol{\mu}} = (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}$, (E.27) now simplifies to

$$\mathcal{T}_{\boldsymbol{\mu}^v} = 2\,\text{diag}\left(\left\{\frac{1-\mu_i}{\mathbb{V}(\text{mc}_{it})}\right\}_{i=1}^{N}\right)\left[\mathbb{COV}\left(\mathbf{mc}_t, \mathbf{mc}_t\right) \odot \left((\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}\mathbf{A}\right)\right], \tag{E.30}$$

where the first term derives from the fixed-point system (3.6) that $\frac{\chi_i}{\theta_i\lambda_i V_i} = 1 - \mu_i$.

For the solution of $\mathcal{T}_{\boldsymbol{\kappa}}$, we follow a similar approach as in Lemma D.4 - D.7, and elaborate on the following proposition.

Proposition E.1. *In the best response function (E.16), the Jacobian matrix $\mathcal{T}_{\boldsymbol{\kappa}}$ in terms of $\boldsymbol{\kappa}$ is given by*

$$\mathcal{T}_{\boldsymbol{\kappa}} = \begin{bmatrix} \frac{\partial \mathcal{T}_1(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} & \frac{\partial \mathcal{T}_2(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} & \cdots & \frac{\partial \mathcal{T}_N(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} \end{bmatrix}'. \tag{E.31}$$

For each sector $i = 1, 2, \ldots N$,

$$\frac{\partial \mathcal{T}_i(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} = \frac{\chi_i}{\theta_i\lambda_i V_i(\boldsymbol{\kappa})^2}\frac{\partial V_i(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'}; \qquad \frac{\partial V_i(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} = 2\boldsymbol{\Sigma}_z\boldsymbol{\Gamma}'\boldsymbol{\Delta}_\mu'\mathbf{e}_i'\mathbf{e}_i\boldsymbol{\Delta}_\mu\boldsymbol{\alpha}. \tag{E.32}$$

Proof. Notice that the defferential of $\boldsymbol{\Gamma} = (-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa})$ is given by

$$d(\boldsymbol{\Gamma}) = \boldsymbol{\alpha} d(\boldsymbol{\kappa}).$$

Under this circumstance, only $\boldsymbol{\kappa}$ is the independent variable in function $\mathcal{T}_i(\boldsymbol{\kappa})$ and $V_i(\boldsymbol{\kappa})$ with given $\boldsymbol{\mu}$. Therefore, using Lemma D.4-D.6, we deduce that

$$\begin{aligned} d\left[V_i(\boldsymbol{\kappa})\right] &= d\left(\mathbf{e}_i\boldsymbol{\Delta}_\mu\boldsymbol{\Gamma}\boldsymbol{\Sigma}_z\boldsymbol{\Gamma}'\boldsymbol{\Delta}_\mu'\mathbf{e}_i'\right) \\ &= \text{tr}\left[\mathbf{e}_i\left(\boldsymbol{\Delta}_\mu d(\boldsymbol{\Gamma})\boldsymbol{\Sigma}_z\boldsymbol{\Gamma}'\boldsymbol{\Delta}_\mu' + \boldsymbol{\Delta}_\mu\boldsymbol{\Gamma}\boldsymbol{\Sigma}_z d(\boldsymbol{\Gamma}')\boldsymbol{\Delta}_\mu'\right)\mathbf{e}_i'\right] \\ &= \text{tr}\left[\mathbf{e}_i\boldsymbol{\Delta}_\mu d(\boldsymbol{\Gamma})\boldsymbol{\Sigma}_z\boldsymbol{\Gamma}'\boldsymbol{\Delta}_\mu'\mathbf{e}_i'\right] + \text{tr}\left[\mathbf{e}_i\boldsymbol{\Delta}_\mu\boldsymbol{\Gamma}\boldsymbol{\Sigma}_z d(\boldsymbol{\Gamma}')\boldsymbol{\Delta}_\mu'\mathbf{e}_i'\right] \\ &= 2\,\text{tr}\left[\mathbf{e}_i\boldsymbol{\Delta}_\mu d(\boldsymbol{\Gamma})\boldsymbol{\Sigma}_z\boldsymbol{\Gamma}'\boldsymbol{\Delta}_\mu'\mathbf{e}_i'\right] \\ &= 2\,\text{tr}\left[\mathbf{e}_i\boldsymbol{\Delta}_\mu\boldsymbol{\alpha} d(\boldsymbol{\kappa})\boldsymbol{\Sigma}_z\boldsymbol{\Gamma}'\boldsymbol{\Delta}_\mu'\mathbf{e}_i'\right] \\ &= 2\,\text{tr}\left[\boldsymbol{\Sigma}_z\boldsymbol{\Gamma}'\boldsymbol{\Delta}_\mu'\mathbf{e}_i'\mathbf{e}_i\boldsymbol{\Delta}_\mu\boldsymbol{\alpha} d(\boldsymbol{\kappa})\right]. \end{aligned} \tag{E.33}$$

By Lemma D.4,

$$d\left[V_i(\boldsymbol{\kappa})\right] = \text{tr}\left[\frac{\partial V_i(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} d\boldsymbol{\kappa}\right], \tag{E.34}$$

where $\frac{\partial V_i(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} = \begin{bmatrix} \frac{\partial V_i(\boldsymbol{\kappa})}{\partial \kappa_1} & \frac{\partial V_i(\boldsymbol{\kappa})}{\partial \kappa_2} & \cdots & \frac{\partial V_i(\boldsymbol{\kappa})}{\partial \kappa_N} \end{bmatrix}'$ serves as a column vector. The juxtaposition of equations (E.33) and (E.34) yields

$$\frac{\partial V_i(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} = 2\boldsymbol{\Sigma}_z \boldsymbol{\Gamma}' \boldsymbol{\Delta}'_\mu \mathbf{e}'_i \mathbf{e}_i \boldsymbol{\Delta}_\mu \boldsymbol{\alpha}; \quad \text{and} \quad \frac{\partial V_i(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} = 2\boldsymbol{\alpha}' \boldsymbol{\Delta}'_\mu \mathbf{e}'_i \mathbf{e}_i \boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z.$$

By definition of $\mathcal{T}_{\boldsymbol{\kappa}}$ in (E.22),

$$\mathcal{T}_{\boldsymbol{\kappa}} = \begin{bmatrix} \frac{\partial \mathcal{T}_1(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} & \frac{\partial \mathcal{T}_2(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} & \cdots & \frac{\partial \mathcal{T}_N(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} \end{bmatrix}'. \tag{E.35}$$

Similar to (D.50), we use the chain rule for matrix derivatives. For each sector $i = 1, 2, \ldots N$,

$$\frac{\partial \mathcal{T}_i(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} = \frac{\chi_i}{\theta_i \lambda_i V_i(\boldsymbol{\kappa})^2} \frac{\partial V_i(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'}; \qquad \frac{\partial V_i(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}'} = 2\boldsymbol{\Sigma}_z \boldsymbol{\Gamma}' \boldsymbol{\Delta}'_\mu \mathbf{e}'_i \mathbf{e}_i \boldsymbol{\Delta}_\mu \boldsymbol{\alpha}, \tag{E.36}$$

which completes the proof of Proposition E.1. □

Given the above Proposition, we transform equations (E.31) and (E.32) using Lemma E.1 as

$$\begin{aligned}
\mathcal{T}_{\boldsymbol{\kappa}} &= \begin{bmatrix} \frac{\partial \mathcal{T}_1(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \\ \frac{\partial \mathcal{T}_2(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \\ \vdots \\ \frac{\partial \mathcal{T}_N(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \end{bmatrix} = \operatorname{diag}\left(\left\{ \frac{\chi_i}{\theta_i \lambda_i V_i^2} \right\}_{i=1}^N \right) \begin{bmatrix} \frac{\partial V_1(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \\ \frac{\partial V_2(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \\ \vdots \\ \frac{\partial V_N(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \end{bmatrix} = 2\operatorname{diag}\left(\left\{ \frac{\chi_i}{\theta_i \lambda_i V_i^2} \right\}_{i=1}^N \right) \begin{bmatrix} \boldsymbol{\alpha}' \boldsymbol{\Delta}'_\mu \mathbf{e}'_1 \mathbf{e}_1 \boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z \\ \boldsymbol{\alpha}' \boldsymbol{\Delta}'_\mu \mathbf{e}'_2 \mathbf{e}_2 \boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z \\ \vdots \\ \boldsymbol{\alpha}' \boldsymbol{\Delta}'_\mu \mathbf{e}'_N \mathbf{e}_N \boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z \end{bmatrix} \\
&= 2\operatorname{diag}\left(\left\{ \frac{\chi_i}{\theta_i \lambda_i V_i^2} \right\}_{i=1}^N \right) \operatorname{diag}\left(\boldsymbol{\Delta}_\mu \boldsymbol{\alpha} \right) \boldsymbol{\Delta}_\mu \boldsymbol{\Gamma} \boldsymbol{\Sigma}_z \\
&= 2\operatorname{diag}\left(\left\{ \frac{1-\mu_i}{V_i} \right\}_{i=1}^N \right) \operatorname{diag}(\boldsymbol{\rho})\, \boldsymbol{\mu}^{-1} \boldsymbol{\phi} \boldsymbol{\Sigma}_z \\
&= 2\operatorname{diag}\left(\left\{ \frac{1-\mu_i}{V_i} \right\}_{i=1}^N \right) \left[\boldsymbol{\mu}^{-1} \boldsymbol{\phi} \mathbb{E}\left[\boldsymbol{z}_t \boldsymbol{z}'_t \right] \odot (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} \boldsymbol{\alpha} \mathbf{1}' \right] \\
&= 2\operatorname{diag}\left(\left\{ \frac{1-\mu_i}{\mathbb{V}(\mathrm{mc}_{it})} \right\}_{i=1}^N \right) \left[\mathbb{COV}\left(\mathbf{mc}_t, \boldsymbol{z}_t \right) \odot (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} \boldsymbol{\alpha} \mathbf{1}' \right]
\end{aligned} \tag{E.37}$$

where we define

$$\boldsymbol{\rho} = (\rho_1, \rho_2, \ldots \rho_N)' = \boldsymbol{\Delta}_{\boldsymbol{\mu}} \boldsymbol{\alpha} = (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1} \boldsymbol{\alpha} \tag{E.38}$$

as in the main text, and implement matrix identity (E.28). The proof of Proposition 4.2 is now complete, according to (E.25), (E.30), and (E.37).

□

E.5 Proof of Lemma 4.3

Proof. We trace the direct and indirect impact of monetary policy $\boldsymbol{\kappa}$ on expected welfare loss $L \equiv \mathbb{E}\left[U_t - U_t^*\right]$ through the lens of functional relations as 4 layers of composite matrix derivatives,

$$L = \mathbb{E}\left[U_t - U_t^*\right] \xrightarrow{\frac{\partial L}{\partial \bar{e}_{it}}} \bar{e}_{it} \xrightarrow{\frac{\partial \bar{e}_{it}}{\partial \mathbf{Q}}} \mathbf{Q} \xrightarrow{\frac{\partial \mathbf{Q}}{\partial \boldsymbol{\mu}^v}} \boldsymbol{\mu}^v \xrightarrow{J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})} \boldsymbol{\kappa}. \tag{E.39}$$

Step 1: Computing Derivatives of Average Pricing Errors $\bar{e}_{it}$.

Recall from the definition of sectoral pricing errors

$$\bar{e}_{it} = \mathbf{e}_i \mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right) \boldsymbol{z}_t, \tag{E.40}$$

with matrix $\mathbf{Q}$ defined as $\mathbf{Q} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu})$.

We start with univariate derivative of a single sector, by chain rule,

$$\frac{d\bar{e}_{it}}{d\kappa_s} = \sum_{r=1}^{N}\sum_{j=1}^{N} \frac{\partial \bar{e}_{it}}{\partial Q_{rj}} \frac{\partial Q_{rj}}{\partial \kappa_s} + \frac{\partial \bar{e}_{it}}{\partial \kappa_s} = \operatorname{tr}\left(J_{\bar{e}_{it}}\left(\mathbf{Q}\right)' J_{\mathbf{Q}}\left(\kappa_s\right)\right) + \frac{\partial \bar{e}_{it}}{\partial \kappa_s} \tag{E.41}$$

where we define two Jocabian matrices of composite derivatives as

$$J_{\bar{e}_{it}}\left(\mathbf{Q}\right) = \begin{bmatrix} \frac{\partial \bar{e}_{it}}{\partial Q_{11}} & \frac{\partial \bar{e}_{it}}{\partial Q_{12}} & \cdots & \frac{\partial \bar{e}_{it}}{\partial Q_{1N}} \\ \frac{\partial \bar{e}_{it}}{\partial Q_{21}} & \frac{\partial \bar{e}_{it}}{\partial Q_{22}} & \cdots & \frac{\partial \bar{e}_{it}}{\partial Q_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \bar{e}_{it}}{\partial Q_{N1}} & \frac{\partial \bar{e}_{it}}{\partial Q_{N2}} & \cdots & \frac{\partial \bar{e}_{it}}{\partial Q_{NN}} \end{bmatrix}; \qquad J_{\mathbf{Q}}\left(\kappa_s\right) = \begin{bmatrix} \frac{\partial Q_{11}}{\partial \kappa_s} & \frac{\partial Q_{12}}{\partial \kappa_s} & \cdots & \frac{\partial Q_{1N}}{\partial \kappa_s} \\ \frac{\partial Q_{21}}{\partial \kappa_s} & \frac{\partial Q_{22}}{\partial \kappa_s} & \cdots & \frac{\partial Q_{2N}}{\partial \kappa_s} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial Q_{N1}}{\partial \kappa_s} & \frac{\partial Q_{N2}}{\partial \kappa_s} & \cdots & \frac{\partial Q_{NN}}{\partial \kappa_s} \end{bmatrix}$$

To solve $J_{\bar{e}_{it}}\left(\mathbf{Q}\right)$, we treat $\boldsymbol{\kappa}$ as a coefficient vector and construct the function of $\bar{e}_{it}\left(\mathbf{Q}\right)$ only with respect to the variable $\mathbf{Q}$. Under such circumstances, the differential of $\bar{e}_{it}\left(\mathbf{Q}\right)$ becomes

$$\begin{aligned} d\left[\bar{e}_{it}\left(\mathbf{Q}\right)\right] &= d\left(\mathbf{e}_i \mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right) \boldsymbol{z}_t\right) \\ &= \operatorname{tr}\left[\mathbf{e}_i d\left(\mathbf{Q}\right)\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right) \boldsymbol{z}_t\right] \\ &= \operatorname{tr}\left[\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right) \boldsymbol{z}_t \mathbf{e}_i d\left(\mathbf{Q}\right)\right], \end{aligned} \tag{E.42}$$

where we apply results from Lemma D.4-D.6. By definition of matrix differential,

$$d\left[\bar{e}_{it}\left(\mathbf{Q}\right)\right] = \operatorname{tr}\left[\frac{\partial\left[\bar{e}_{it}\left(\mathbf{Q}\right)\right]}{\partial \mathbf{Q}'} d\left(\mathbf{Q}\right)\right] = \operatorname{tr}\left[J_{\bar{e}_{it}}\left(\mathbf{Q}\right)' d\left(\mathbf{Q}\right)\right]. \tag{E.43}$$

Since $d\left(\mathbf{Q}\right)$ is a matrix of differentials for variable $\mathbf{Q}$, in order to equalize (E.42) and (E.43), we must have

$$J_{\bar{e}_{it}}\left(\mathbf{Q}\right)' = \frac{\partial\left[\bar{e}_{it}\left(\mathbf{Q}\right)\right]}{\partial \mathbf{Q}'} = \left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right) \boldsymbol{z}_t \mathbf{e}_i. \tag{E.44}$$

By inspection, the othornomal unit row vector $\mathbf{e}_i$ determines that the elements of $J_{\bar{e}_{it}}(\mathbf{Q})'$ are 0 except for the ith column:

$$J_{\bar{e}_{it}}(\mathbf{Q})' = \begin{bmatrix} 0 & \cdots & \frac{\partial \bar{e}_{it}}{\partial Q_{i1}} & \cdots & 0 \\ 0 & \cdots & \frac{\partial \bar{e}_{it}}{\partial Q_{i2}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \frac{\partial \bar{e}_{it}}{\partial Q_{iN}} & \cdots & 0 \end{bmatrix}; \qquad \left[J_{\bar{e}_{it}}(\mathbf{Q})'\right]_{(:,i)} = (\mathbf{L} - \mathbf{1}\boldsymbol{\kappa})\, \boldsymbol{z}_t,$$

and $\left[J_{\bar{e}_{it}}(\mathbf{Q})\right]_{(i,:)} = \boldsymbol{z}_t'(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa})'$, It follows that (E.41) simplifies to

$$\frac{d\bar{e}_{it}}{d\kappa_s} = \left[J_{\bar{e}_{it}}(\mathbf{Q})\right]_{(i,:)} \left[J_{\mathbf{Q}}(\kappa_s)'\right]_{(:,i)} + \frac{\partial \bar{e}_{it}}{\partial \kappa_s} = \boldsymbol{z}_t'(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa})' \left[J_{\mathbf{Q}}(\kappa_s)'\right]_{(:,i)} + \frac{\partial \bar{e}_{it}}{\partial \kappa_s}, \tag{E.45}$$

where the first equality follows from definition of trace operator, and

$$\left[J_{\mathbf{Q}}(\kappa_s)'\right]_{(:,i)} = \begin{bmatrix} \frac{\partial Q_{i1}}{\partial \kappa_s} & \frac{\partial Q_{i2}}{\partial \kappa_s} & \cdots & \frac{\partial Q_{iN}}{\partial \kappa_s} \end{bmatrix}'.$$

Using the same matrix calculus techniques from (E.42) - (E.44),

$$\frac{\partial \bar{e}_{it}}{\partial \boldsymbol{\kappa}} = -\boldsymbol{z}_t \mathbf{e}_i \mathbf{Q1} = -[\mathbf{Q1}]_{(i)} \boldsymbol{z}_t; \qquad \frac{\partial \bar{e}_{it}}{\partial \kappa_s} = -[\mathbf{Q1}]_{(i)} z_{st}; \qquad i, s = 1, 2, ...N,$$

where $[\mathbf{Q1}]_{(i)} = \mathbf{e}_i \mathbf{Q1}$ is scalar. Therefore, the univariate derivative of average pricing errors in (E.45) admits the form,

$$\frac{d\bar{e}_{it}}{d\kappa_s} = \boldsymbol{z}_t'(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa})' \left[J_{\mathbf{Q}}(\kappa_s)'\right]_{(:,i)} - [\mathbf{Q1}]_{(i)} z_{st}. \tag{E.46}$$

Next, we characterize the key derivative $\left[J_{\mathbf{Q}}(\kappa_s)'\right]_{(:,i)}$ in (E.41). By definition and chain rule,

$$\left[J_{\mathbf{Q}}(\kappa_s)'\right]_{(:,i)} = \begin{bmatrix} \frac{\partial Q_{i1}}{\partial \kappa_s} \\ \frac{\partial Q_{i2}}{\partial \kappa_s} \\ \vdots \\ \frac{\partial Q_{iN}}{\partial \kappa_s} \end{bmatrix} = \begin{bmatrix} \frac{\partial Q_{i1}}{\partial (\boldsymbol{\mu}^v)'} \\ \frac{\partial Q_{i2}}{\partial (\boldsymbol{\mu}^v)'} \\ \vdots \\ \frac{\partial Q_{iN}}{\partial (\boldsymbol{\mu}^v)'} \end{bmatrix} \left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}. \tag{E.47}$$

In particular,

$$\frac{\partial Q_{ij}}{\partial (\boldsymbol{\mu}^v)'} = \begin{bmatrix} \frac{\partial Q_{ij}}{\partial \mu_1} & \frac{\partial Q_{ij}}{\partial \mu_2} & \cdots & \frac{\partial Q_{ij}}{\partial \mu_N} \end{bmatrix}; \quad \forall i, j = 1, 2, \ldots.N,$$

and $J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa}) = \begin{bmatrix} \frac{\partial \mu_1}{\partial \kappa_1} & \frac{\partial \mu_1}{\partial \kappa_2} & \cdots & \frac{\partial \mu_1}{\partial \kappa_N} \\ \frac{\partial \mu_2}{\partial \kappa_1} & \frac{\partial \mu_2}{\partial \kappa_2} & \cdots & \frac{\partial \mu_2}{\partial \kappa_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mu_N}{\partial \kappa_1} & \frac{\partial \mu_2}{\partial \kappa_2} & \cdots & \frac{\partial \mu_N}{\partial \kappa_N} \end{bmatrix}$ is constructed in (E.23), following the expression in Proposition 4.2.

To solve for $\frac{\partial Q_{ij}}{\partial(\boldsymbol{\mu}^v)'}$ in (E.47), we note that the differential of $\mathbf{Q} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu})$ is given by

$$\begin{aligned} d\mathbf{Q} &= d\,(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu}) + (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\,d\,(\mathbf{I} - \boldsymbol{\mu}) \\ &= -(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\,d\,(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu}) + (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\,d\,(\mathbf{I} - \boldsymbol{\mu}) \\ &= (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\,d\boldsymbol{\mu}\mathbf{A}\,(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu}) - (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\,d\boldsymbol{\mu} \\ &= (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\,d\boldsymbol{\mu}\left[\mathbf{A}\,(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu}) - \mathbf{I}\right] \\ &= (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\,d\boldsymbol{\mu}\,(\mathbf{A}\mathbf{Q} - \mathbf{I})\,. \end{aligned} \tag{E.48}$$

For each element of $d\mathbf{Q} \in \mathbb{R}^{N\times N}$,

$$[d\mathbf{Q}]_{(i,j)} = \sum_{l=1}^{N}\left[(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\right]_{(i,l)}[(\mathbf{A}\mathbf{Q} - \mathbf{I})]_{(l,j)}\,d\mu_l; \qquad \forall i, j = 1, 2, ...N,$$

where I use the diagonal property of $\boldsymbol{\mu}$. Hence,

$$\frac{\partial\,[Q]_{(i,j)}}{\partial\mu_l} = \frac{[dQ]_{(i,j)}}{d\mu_l} = \left[(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\right]_{(i,l)}[(\mathbf{A}\mathbf{Q} - \mathbf{I})]_{(l,j)}\,; \qquad \forall i, j, l = 1, 2, ...N.$$

Stack the above equations in vector form,

$$\frac{\partial\,[\mathbf{Q}]_{(i,j)}}{\partial\boldsymbol{\mu}^v} = \left\{\left[(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\right]_{(i,:)}\right\}' \odot [(\mathbf{A}\mathbf{Q} - \mathbf{I})]_{(:,j)}\,; \qquad \forall i, j = 1, 2, ...N,$$

which can be transposed to

$$\frac{\partial\,[\mathbf{Q}]_{(i,j)}}{\partial\,(\boldsymbol{\mu}^v)'} = \left(\frac{\partial\,[\mathbf{Q}]_{(i,j)}}{\partial\boldsymbol{\mu}^v}\right)' = \left[(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\right]_{(i,:)} \odot \left[(\mathbf{A}\mathbf{Q} - \mathbf{I})'\right]_{(j,:)}\,; \qquad \forall i, j = 1, 2, ...N, \tag{E.49}$$

If we use the following matrix identities,

$$\mathbf{A}\mathbf{Q} - \mathbf{I} = -\boldsymbol{\mu}^{-1}(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}(\mathbf{I} - \mathbf{A}) = -(\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}(\mathbf{I} - \mathbf{A}),$$

we can stack the vector in (E.49) into matrix form as

$$\begin{bmatrix} \frac{\partial Q_{i1}}{\partial(\boldsymbol{\mu}^v)'} \\ \frac{\partial Q_{i2}}{\partial(\boldsymbol{\mu}^v)'} \\ \vdots \\ \frac{\partial Q_{iN}}{\partial(\boldsymbol{\mu}^v)'} \end{bmatrix} = (\mathbf{A}\mathbf{Q} - \mathbf{I})'\,\mathrm{diag}\left(\left[(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\right]_{(i,:)}\right) = -\left[(\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}(\mathbf{I} - \mathbf{A})\right]'\,\mathrm{diag}\left(\left[(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\right]_{(i,:)}\right).$$

The first equality follows from the property of Hadamard product. As such, we obtain a representation of (E.47),

$$\left[J_{\mathbf{Q}}\left(\kappa_s\right)'\right]_{(:,i)} = -\left[(\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1}(\mathbf{I}-\mathbf{A})\right]' \operatorname{diag}\left(\left[(\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}\right]_{(i,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)} \tag{E.50}$$

Before we move on to the next step, we express the univariate derivative by substituting (E.50) into (E.46),

$$\begin{aligned}\frac{d\bar{e}_{it}}{d\kappa_s} &= \left[(\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1}(\mathbf{I}-\mathbf{A})\left(\mathbf{1}\boldsymbol{\kappa}-\mathbf{L}\right)\boldsymbol{z}_t\right]' \operatorname{diag}\left(\left[(\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}\right]_{(i,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)} - [\mathbf{Q1}]_{(i)}z_{st} \\ &= \boldsymbol{z}_t'\left[\boldsymbol{\mu}^{-1}(\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}(-\mathbf{I}+\boldsymbol{\alpha\kappa})\right]' \operatorname{diag}\left(\left[(\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}\right]_{(i,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)} - [\mathbf{Q1}]_{(i)}z_{st} \\ &= \boldsymbol{z}_t'\boldsymbol{\phi}'\boldsymbol{\mu}^{-1}\operatorname{diag}\left([\mathbf{H}]_{(i,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)} - [\mathbf{Q1}]_{(i)}z_{st}.\end{aligned} \tag{E.51}$$

where the second equality follows from the matrix identities (D.36) and $(\mathbf{1}\boldsymbol{\kappa}-\mathbf{L}) = \mathbf{L}(-\mathbf{I}+\boldsymbol{\alpha\kappa})$, and the last equality from the definition of $\boldsymbol{\phi} = (\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}(-\mathbf{I}+\boldsymbol{\alpha\kappa})$ and $\mathbf{H} = (\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}$. Equation (E.51) measures the total marginal impact of wage rule $\boldsymbol{\kappa}$ on the average pricing errors $\bar{e}_{it}$. We now turn to the step 2.

Step 2: Computing Derivatives of Expected Welfare Loss.

The second-order approximation of the expected welfare loss (4.1) can be expressed as

$$\begin{aligned}L &= \Delta U_t^{within} + \Delta U_t^{OG} + \Delta U_t^{across} \\ &= \frac{1}{2}\mathbb{E}\left[\sum_{i=1}^{N}\chi_i\mu_i + \frac{1}{(\gamma+1/\eta)}\left(\sum_{j=1}^{n}\beta_j\bar{e}_{jt}\right)^2 + \sum_{i=1}^{N}\lambda_i\sum_{j=1}^{n}a_{ij}\bar{e}_{jt}^2 - \sum_{i=1}^{N}\lambda_i\left(\sum_{j=1}^{n}a_{ij}\bar{e}_{jt}\right)^2 + \sum_{j=1}^{n}\beta_j\bar{e}_{jt}^2 - \left(\sum_{j=1}^{n}\beta_j\bar{e}_{jt}\right)^2\right] \\ &= \frac{1}{2}\mathbb{E}\left[\sum_{i=1}^{N}\chi_i\mu_i + \left(\frac{1}{(\gamma+1/\eta)}-1\right)\left(\sum_{j=1}^{n}\beta_j\bar{e}_{jt}\right)^2 + \sum_{i=1}^{N}\lambda_i\sum_{j=1}^{n}a_{ij}\bar{e}_{jt}^2 - \sum_{i=1}^{N}\lambda_i\left(\sum_{j=1}^{n}a_{ij}\bar{e}_{jt}\right)^2 + \sum_{j=1}^{n}\beta_j\bar{e}_{jt}^2\right]\end{aligned} \tag{E.52}$$

with first-order condition with respect to $\kappa_s, s = 1, 2, ...N$ given by

$$\underbrace{\frac{1}{2}\sum_{i=1}^{N}\mathbb{E}\left[\chi_i\frac{d\mu_i}{d\kappa_s}\right]}_{\text{Term (1)}} + \left(\frac{1}{(\gamma+1/\eta)}-1\right)\underbrace{\sum_{i=1}^{N}\sum_{j=1}^{N}\beta_i\beta_j\mathbb{E}\left[\bar{e}_{it}\frac{d\bar{e}_{jt}}{d\kappa_s}\right]}_{\text{Term (2)}} + \underbrace{\sum_{i=1}^{N}\lambda_i\sum_{j=1}^{N}a_{ij}\mathbb{E}\left[\bar{e}_{jt}\frac{d\bar{e}_{jt}}{d\kappa_s}\right]}_{\text{Term (3)}} - \underbrace{\sum_{i=1}^{N}\lambda_i\sum_{j=1}^{N}\sum_{r=1}^{N}a_{ij}a_{ir}\mathbb{E}\left[\bar{e}_{rt}\frac{d\bar{e}_{jt}}{d\kappa_s}\right]}_{\text{Term (4)}} + \underbrace{\sum_{j=1}^{N}\beta_j\mathbb{E}\left[\bar{e}_{jt}\frac{d\bar{e}_{jt}}{d\kappa_s}\right]}_{\text{Term (5)}} = 0 \tag{E.53}$$

Next, we compute term-by-term the expected derivatives in (E.53) using results we have derived above. In deriving these matrix expressions, we need several matrix identities involving Hadamard product and diagonal matrices, summarized in the following lemma.

Lemma E.2. *Let $\mathbf{X}_a$ and $\mathbf{X}_b$ denote two N-dimensional column vectors, then*

$$\mathbf{X}_a' \operatorname{diag}(X_b) = \mathbf{X}_a' \odot \mathbf{X}_b'$$

Let $\mathbf{X}_A$ and $\mathbf{X}_B$ denote two $N \times N$ matrices, with elements of $\mathbf{X}_A$ denoted by $[\mathbf{X}_A]_{ij} = x_{ij}^A$. Then

$$\sum_{j=1}^{N} x_{ij}^A \mathbf{e}_j \mathbf{X}_B = [\mathbf{X}_A]_{(i,:)} \mathbf{X}_B$$

Combining these two results together with the definition of Hadamard product,

$$\begin{bmatrix} \mathbf{e}_1 \mathbf{X}_A \ \mathrm{diag}\left([\mathbf{X}_B]_{(1,:)}\right) \\ \mathbf{e}_2 \mathbf{X}_A \ \mathrm{diag}\left([\mathbf{X}_B]_{(2,:)}\right) \\ \vdots \\ \mathbf{e}_N \mathbf{X}_A \ \mathrm{diag}\left([\mathbf{X}_B]_{(3,:)}\right) \end{bmatrix} = \begin{bmatrix} [\mathbf{X}_A]_{(1,:)} \odot [\mathbf{X}_B]_{(1,:)} \\ [\mathbf{X}_A]_{(2,:)} \odot [\mathbf{X}_B]_{(2,:)} \\ \vdots \\ [\mathbf{X}_A]_{(N,:)} \odot [\mathbf{X}_B]_{(N,:)} \end{bmatrix} = \mathbf{X}_A \odot \mathbf{X}_B$$

Finally, suppose $\mathbf{X}_C$ is a $N \times N$ matrix,

$$\begin{bmatrix} \left([\mathbf{X}_C]_{(1,:)} \mathbf{X}_A\right) \odot \left([\mathbf{X}_C]_{(1,:)} \mathbf{X}_B\right) \\ \left([\mathbf{X}_C]_{(2,:)} \mathbf{X}_A\right) \odot \left([\mathbf{X}_C]_{(2,:)} \mathbf{X}_B\right) \\ \vdots \\ \left([\mathbf{X}_C]_{(N,:)} \mathbf{X}_A\right) \odot \left([\mathbf{X}_C]_{(N,:)} \mathbf{X}_B\right) \end{bmatrix} = \begin{bmatrix} \left([\mathbf{X}_C]_{(1,:)} \mathbf{X}_A\right) \\ \left([\mathbf{X}_C]_{(2,:)} \mathbf{X}_A\right) \\ \vdots \\ \left([\mathbf{X}_C]_{(N,:)} \mathbf{X}_A\right) \end{bmatrix} \odot \begin{bmatrix} \left([\mathbf{X}_C]_{(1,:)} \mathbf{X}_B\right) \\ \left([\mathbf{X}_C]_{(2,:)} \mathbf{X}_B\right) \\ \vdots \\ \left([\mathbf{X}_C]_{(N,:)} \mathbf{X}_B\right) \end{bmatrix} = \left(\mathbf{X}_C \mathbf{X}_A\right) \odot \left(\mathbf{X}_C \mathbf{X}_B\right)$$

The proof of Lemma E.2 is straightforward and hence omitted. Using this lemma, we carry out the derivation of (E.53).

Solution of Term (1). The first term of (E.53) follows

$$\text{Term (1)} = \sum_{i=1}^{N} \mathbb{E}\left[\chi_i \frac{d\mu_i}{d\kappa_s}\right] = \sum_{i=1}^{N} \chi_i \frac{d\mu_i}{d\kappa_s} = \boldsymbol{\chi}' \frac{d\boldsymbol{\mu}^v}{d\kappa_s} = \boldsymbol{\chi}' \left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)} \tag{E.54}$$

where $J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})$ is given in the proof of Proposition 4.2.

Solution of Term (2). To begin with, we notice that a matrix identity holds as

$$\mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right) = -\mathbf{Q}\mathbf{L}(-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa}) = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\left[\boldsymbol{\mu}(\mathbf{I} - \mathbf{A}) - (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})\right]\mathbf{L}\boldsymbol{\mu}^{-1}(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})\boldsymbol{\phi} = \left[\mathbf{I} - \mathbf{L}(\boldsymbol{\mu}^{-1} - \mathbf{A})\right]\boldsymbol{\phi} = \mathbf{L}(\mathbf{I} - \boldsymbol{\mu}^{-1})\boldsymbol{\phi} \tag{E.55}$$

where I use the matrix identity $\mathbf{L}\boldsymbol{\alpha} = \mathbf{1}$, and adopt the definition of $\boldsymbol{\phi} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}(-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa})$ and $\mathbf{Q} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu})$. With equation (E.55) and the definition of $\mathbb{V}\left(\mathbf{mc}_t\right) = \boldsymbol{\mu}^{-1}\boldsymbol{\phi}\boldsymbol{\Sigma}_z\boldsymbol{\phi}'\boldsymbol{\mu}^{-1}$ in (E.29), we obtain

$$\mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{\Sigma}_z\boldsymbol{\phi}'\boldsymbol{\mu}^{-1} = \mathbf{L}(\boldsymbol{\mu} - \mathbf{I})\mathbb{V}\left(\mathbf{mc}_t\right). \tag{E.56}$$

Next, we use equations (E.40) and (E.51) to derive

$$\begin{aligned}
\mathbb{E}\left[\bar{e}_{it}\frac{d\bar{e}_{jt}}{d\kappa_s}\right] &= \mathbb{E}\left\{\mathbf{e}_i\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{z}_t\left[\boldsymbol{z}_t'\boldsymbol{\phi}'\boldsymbol{\mu}^{-1}\operatorname{diag}\left([\mathbf{H}]_{(j,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-[\mathbf{Q1}]_{(j)}z_{st}\right]\right\}\\
&= \mathbf{e}_i\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{\Sigma}_z\boldsymbol{\phi}'\boldsymbol{\mu}^{-1}\operatorname{diag}\left([\mathbf{H}]_{(j,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-[\mathbf{Q1}]_{(j)}\mathbf{e}_i\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)[\boldsymbol{\Sigma}_z]_{(:,s)}\\
&= \mathbf{e}_i\mathbf{L}(\boldsymbol{\mu}-\mathbf{I})\Sigma_{\mathbf{p}_t^{\Delta}}\operatorname{diag}\left([\mathbf{H}]_{(j,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-[\mathbf{Q1}]_{(j)}\mathbf{e}_i\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)[\boldsymbol{\Sigma}_z]_{(:,s)},
\end{aligned}$$

where the second identity follows from the definition that $\mathbb{E}\left[\boldsymbol{z}_t\boldsymbol{z}_t'\right]=\boldsymbol{\Sigma}_z$ and $\mathbb{E}\left[\boldsymbol{z}_t z_s\right]=[\boldsymbol{\Sigma}_z]_{(:,s)}$, and the last identity follows the matrix identity of (E.56).

The second term of (E.53) now follows

$$\begin{aligned}
\text{Term (2)} &= \sum_{i=1}^{N}\sum_{j=1}^{N}\beta_i\beta_j\mathbb{E}\left[\bar{e}_{it}\frac{d\bar{e}_{jt}}{d\kappa_s}\right]\\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}\beta_i\beta_j\left[\mathbf{e}_i\mathbf{L}(\boldsymbol{\mu}-\mathbf{I})\mathbb{V}\left(\mathbf{mc}_t\right)\operatorname{diag}\left([\mathbf{H}]_{(j,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-[\mathbf{Q1}]_{(j)}\mathbf{e}_i\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)[\boldsymbol{\Sigma}_z]_{(:,s)}\right]\\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\beta_i\mathbf{e}_i\mathbf{L}(\boldsymbol{\mu}-\mathbf{I})\mathbb{V}\left(\mathbf{mc}_t\right)\right)\left[\beta_j\operatorname{diag}\left([\mathbf{H}]_{(j,:)}\right)\right]\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-\sum_{j=1}^{N}\beta_j[\mathbf{Q1}]_{(j)}\sum_{i=1}^{N}\beta_i\mathbf{e}_i\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)[\boldsymbol{\Sigma}_z]_{(:,s)}\\
&= \left[\left(\boldsymbol{\beta}'\mathbf{L}\left(\boldsymbol{\mu}-\mathbf{I}\right)\mathbb{V}\left(\mathbf{mc}_t\right)\right)\odot\left(\boldsymbol{\beta}'\mathbf{H}\right)\right]\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-\left(\boldsymbol{\beta}'\mathbf{Q1}\right)\boldsymbol{\beta}'\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)[\boldsymbol{\Sigma}_z]_{(:,s)}\\
&= \mathbf{r}^{\beta'}\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-\left(\boldsymbol{\beta}'\mathbf{Q1}\right)\boldsymbol{\beta}'\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)[\boldsymbol{\Sigma}_z]_{(:,s)},
\end{aligned}\tag{E.57}$$

where we exploit the property of Hadamard product in Lemma E.2 and define

$$\mathbf{r}^{\beta'}=\left[\left(\boldsymbol{\beta}'\mathbf{L}\left(\boldsymbol{\mu}-\mathbf{I}\right)\mathbb{V}\left(\mathbf{mc}_t\right)\right)\odot\left(\boldsymbol{\beta}'\mathbf{H}\right)\right]\tag{E.58}$$

where $\mathbf{H}=(\mathbf{I}-\boldsymbol{\mu}\mathbf{A})^{-1}$.

Solution of Term (3). The third term in (E.53) follows

$$\begin{aligned}
\text{Term (3)} &= \sum_{i=1}^{N}\lambda_i\sum_{j=1}^{N}a_{ij}\mathbb{E}\left[\bar{e}_{jt}\frac{d\bar{e}_{jt}}{d\kappa_s}\right]\\
&= \sum_{i=1}^{N}\lambda_i\sum_{j=1}^{N}a_{ij}\left[\mathbf{e}_j\mathbf{L}(\boldsymbol{\mu}-\mathbf{I})\Sigma_{\mathbf{p}_t^{\Delta}}\operatorname{diag}\left([\mathbf{H}]_{(j,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-[\mathbf{Q1}]_{(j)}\mathbf{e}_j\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)[\boldsymbol{\Sigma}_z]_{(:,s)}\right]\\
&= \sum_{i=1}^{N}\lambda_i\left[[\mathbf{A}]_{(i,:)}\left[\left(\mathbf{L}\left(\boldsymbol{\mu}-\mathbf{I}\right)\Sigma_{\mathbf{p}_t^{\Delta}}\right)\odot\mathbf{H}\right]\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-[\mathbf{A}]_{(i,:)}\operatorname{diag}\left(\mathbf{Q1}\right)\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)[\boldsymbol{\Sigma}_z]_{(:,s)}\right]\\
&= \boldsymbol{\lambda}'\mathbf{A}\mathbf{R}^{\mathbf{I}}\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-\boldsymbol{\lambda}'\mathbf{A}\operatorname{diag}\left(\mathbf{Q1}\right)\mathbf{Q}\left(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa}\right)[\boldsymbol{\Sigma}_z]_{(:,s)}
\end{aligned}\tag{E.59}$$

where we define $\mathbf{R}^I = [(\mathbf{L}(\boldsymbol{\mu}-\mathbf{I})\mathbb{V}(\mathbf{mc}_t))\odot\mathbf{H}]$.

Solution of Term (4). Similarly, the fourth term of (E.53) follows

$$\begin{aligned}
\text{Term (4)} &= \sum_{i=1}^{N}\lambda_i\sum_{j=1}^{N}\sum_{r=1}^{N}a_{ij}a_{ir}\mathbb{E}\left[\bar{e}_{rt}\frac{d\bar{e}_{jt}}{d\kappa_s}\right]\\
&= \sigma_s^2\sum_{i=1}^{N}\lambda_i\sum_{j=1}^{N}\sum_{r=1}^{N}a_{ij}a_{ir}\left[\mathbf{e}_r\mathbf{L}(\boldsymbol{\mu}-\mathbf{I})\mathbb{V}(\mathbf{mc}_t)\,\text{diag}\left([\mathbf{H}]_{(j,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-[\mathbf{Q1}]_{(j)}\mathbf{e}_r\mathbf{Q}(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})[\boldsymbol{\Sigma}_z]_{(:,s)}\right]\\
&= \sum_{i=1}^{N}\lambda_i\left\{\sum_{r=1}^{N}\sum_{j=1}^{N}(a_{ir}e_r\mathbf{L}(\boldsymbol{\mu}-\mathbf{I})\mathbb{V}(\mathbf{mc}_t))\left[a_{ij}\ \text{diag}\left([\mathbf{H}]_{(j,:)}\right)\right]\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-\sum_{j=1}^{N}a_{ij}[\mathbf{Q1}]_{(j)}\sum_{r=1}^{N}a_{ir}e_r\mathbf{Q}(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})[\boldsymbol{\Sigma}_z]_{(:,s)}\right\}\\
&= \sum_{i=1}^{N}\lambda_i\left\{\left[\left([\mathbf{A}]_{(i,:)}\mathbf{L}(\boldsymbol{\mu}-\mathbf{I})\mathbb{V}(\mathbf{mc}_t)\right)\odot\left([\mathbf{A}]_{(i,:)}\mathbf{H}\right)\right]\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-\left([\mathbf{A}]_{(i,:)}\mathbf{Q1}\right)\left([\mathbf{A}]_{(i,:)}\mathbf{Q}\right)(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})[\boldsymbol{\Sigma}_z]_{(:,s)}\right\}\\
&= \boldsymbol{\lambda}'[(\mathbf{AL}(\boldsymbol{\mu}-\mathbf{I})\mathbb{V}(\mathbf{mc}_t))\odot(\mathbf{AH})]\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-\boldsymbol{\lambda}'\,\text{diag}(\mathbf{AQ1})\,\mathbf{AQ}(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})[\boldsymbol{\Sigma}_z]_{(:,s)}\\
&= \boldsymbol{\lambda}'\mathbf{R}^{\mathbf{A}}\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-\boldsymbol{\lambda}'\,\text{diag}(\mathbf{AQ1})\,\mathbf{AQ}(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})[\boldsymbol{\Sigma}_z]_{(:,s)}
\end{aligned} \tag{E.60}$$

where we define $\mathbf{R}^A = [(\mathbf{AL}(\boldsymbol{\mu}-\mathbf{I})\mathbb{V}(\mathbf{mc}_t))\odot(\mathbf{AH})]$.

Solution of Term (5). Finally, the fifth term of (E.53) follows

$$\begin{aligned}
\text{Term (5)} &= \sum_{j=1}^{N}\beta_j\mathbb{E}\left[\bar{e}_{jt}\frac{d\bar{e}_{jt}}{d\kappa_s}\right]\\
&= \sum_{j=1}^{N}\beta_j\left[\mathbf{e}_j\mathbf{L}(\boldsymbol{\mu}-\mathbf{I})\mathbb{V}(\mathbf{mc}_t)\,\text{diag}\left([\mathbf{H}]_{(j,:)}\right)\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-[\mathbf{Q1}]_{(j)}\mathbf{e}_j\mathbf{Q}(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})[\boldsymbol{\Sigma}_z]_{(:,s)}\right]\\
&= \boldsymbol{\beta}'[(\mathbf{L}(\boldsymbol{\mu}-\mathbf{I})\mathbb{V}(\mathbf{mc}_t))\odot\mathbf{H}]\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-\boldsymbol{\beta}'\,\text{diag}(\mathbf{Q1})\,\mathbf{Q}(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})[\boldsymbol{\Sigma}_z]_{(:,s)}\\
&= \boldsymbol{\beta}'\mathbf{R}^{\mathbf{I}}\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)}-\boldsymbol{\beta}'\,\text{diag}(\mathbf{Q1})\,\mathbf{Q}(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})[\boldsymbol{\Sigma}_z]_{(:,s)}
\end{aligned} \tag{E.61}$$

To summarize, we have obtained the derivative expressions of the five terms in the first-order condition (E.53): (E.54), (E.57), (E.59), (E.60) and (E.61). We collect these five expressions according to the three welfare terms expressed in (E.52), and arrange them in row for all $\kappa_s, s = 1, 2, ...N$. The functional relations between attention and welfare loss is characterized in (E.39) . Therefore, using (E.51) and (E.53), it is straightforward to show that the welfare exposure to sectoral attention change is

$$\frac{\partial L}{\partial(\boldsymbol{\mu}^v)'} = \boldsymbol{r}^o + \boldsymbol{r}^c + \boldsymbol{r}^d.$$

In particular, the three row vectors are all associated with the Jocabian matrix $J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})$,

$$\boldsymbol{r}^o = \frac{1}{\gamma + 1/\eta}\mathbf{r}^{\beta'} = \frac{1}{\gamma + 1/\eta}\left[(\boldsymbol{\beta}'\mathbf{M}) \odot (\boldsymbol{\beta}'\mathbf{H})\right], \tag{E.62}$$

$$\boldsymbol{r}^c = \boldsymbol{\lambda}'\mathbf{R}^{\mathbf{I}} - \boldsymbol{\lambda}'\mathbf{R}^{\mathbf{A}} - \boldsymbol{r}^o = \boldsymbol{\lambda}'\left[\mathbf{M} \odot \mathbf{H}\right] - \boldsymbol{\lambda}'\left[(\mathbf{AM}) \odot (\mathbf{AH})\right] - \boldsymbol{r}^o, \tag{E.63}$$

$$\boldsymbol{r}^d = \frac{1}{2}\chi'. \tag{E.64}$$

where $\mathbf{M} = \mathbf{L}(\boldsymbol{\mu} - \mathbf{I})\mathbb{V}(\mathbf{mc}_t)$ and $\mathbf{H} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}$. Note that we have combined (E.59) and (E.61) as

$$\boldsymbol{\lambda}'\mathbf{A}\mathbf{R}^{\mathbf{I}} + \boldsymbol{\beta}'\mathbf{R}^{\mathbf{I}} = \boldsymbol{\lambda}'\mathbf{R}^{\mathbf{I}}$$

and $\boldsymbol{\lambda}'\mathbf{A} = \boldsymbol{\lambda}' - \boldsymbol{\beta}'$. The proof is now complete.

□

E.6 Proof of Proposition 4.3

Proof. We assume that an interior optimal monetary policy exists such that $\mathbf{0} << \boldsymbol{\mu}^v << \mathbf{1}$. We proceed with the proof in three steps.

Step 1: FOC of Optimal Wage Rule as Matrix Fixed-Point.
In contrast to exogenous-information models, the vector-valued FOC equation (4.2) for $\boldsymbol{\kappa}$ alone cannot pin down the optimal monetary policy. To see this point, we note that Definition 4 suggests the optimal policy relies on sectoral attentions $\boldsymbol{\mu}$, but such choice of optimal policy in turn determines the equilibrium distributions of attentions via (E.7). Motivated by this observation, we characterize the solution of optimal policy in our RI equilibrium as a joint fixed-point between policy and attention.

Proposition E.2. *If an interior optimal monetary policy exists such that $\mathbf{0} << \boldsymbol{\mu}^v << \mathbf{1}$, it admits the following characterization:*
(i) Conditional on equilibrium sectoral attention allocation $\boldsymbol{\mu}$ and the covariance matrix of the equilibrium marginal costs $\mathbb{V}(\boldsymbol{mc}_t)$, the optimal monetary policy (wage) rule $\boldsymbol{\kappa}$ is determined by the fixed-point equation:

$$\boldsymbol{\kappa} = \frac{\boldsymbol{\Phi}(\boldsymbol{\kappa})}{\boldsymbol{\Phi}(\boldsymbol{\kappa})\boldsymbol{\alpha}}. \tag{E.65}$$

The $1 \times N$ vector $\boldsymbol{\Phi}(\boldsymbol{\kappa})$ is given by

$$\boldsymbol{\Phi} = \left\{\left(\frac{1}{(\gamma + 1/\eta)} - 1\right)\left(\boldsymbol{\beta}^T\mathbf{Q}\mathbf{1}\right)\boldsymbol{\beta}^T + \boldsymbol{\lambda}^T\Big(\operatorname{diag}(\mathbf{Q}\mathbf{1}) - \operatorname{diag}(\mathbf{A}\mathbf{Q}\mathbf{1})\,\mathbf{A}\Big)\right\}\mathbf{QL} + \left\{\boldsymbol{r}^o + \boldsymbol{r}^c + \boldsymbol{r}^d\right\}\mathcal{R}\operatorname{diag}\left(\left\{\frac{2(1-\mu_i)\rho_i}{\mu_i V_i}\right\}_{i=1}^N\right)(\mathbf{I} - \mathbf{Q})\mathbf{L} \tag{E.66}$$

where we define $\mathcal{R} \equiv \left[I - \mathcal{T}_{\boldsymbol{\mu}^v}\right]^{-1}$, which measures the general equilibrium feedback due to strategic complementarity in information acquisition.

(ii). Conditional on optimal policy rule $\boldsymbol{\kappa}$ determined by (E.65), the equilibrium attention $\boldsymbol{\mu}$ and the covariance structure of marginal costs, $\mathbb{V}(\boldsymbol{mc}_t)$, are endogenously determined via the fixed-point system (E.7), which implicitly define the mapping from optimal $\boldsymbol{\kappa}$ to $\boldsymbol{\mu}$ as

$$\boldsymbol{\mu} = \mathcal{G}(\boldsymbol{\kappa})$$

Therefore, the optimal wage rule $\boldsymbol{\kappa}$ and the endogenous degree of nominal rigidities $\boldsymbol{\mu}$ are determined jointly by fixed-point system (E.7) and (E.65).

Proof. Using expressions in (E.54), (E.57), (E.59), (E.60) and (E.61), combining terms, and empolying the matrix identity $\boldsymbol{\lambda}'A = \boldsymbol{\lambda}' - \boldsymbol{\beta}'$ from (D.5), the FOC (E.53) now becomes

$$\begin{aligned}&\left\{\frac{1}{2}\boldsymbol{\chi}' + \left(\frac{1}{(\gamma+1/\eta)} - 1\right)\boldsymbol{r}^{\beta'} + \boldsymbol{\lambda}'\left(\mathbf{R}^{\mathbf{I}} - \mathbf{R}^{\mathbf{A}}\right)\right\}\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)} + \\ &\left\{-\left(\frac{1}{(\gamma+1/\eta)} - 1\right)(\boldsymbol{\beta}'\mathbf{Q1})\,\boldsymbol{\beta}' - \boldsymbol{\lambda}'\Big(\operatorname{diag}(\mathbf{Q1}) - \operatorname{diag}(\mathbf{AQ1})\,\mathbf{A}\Big)\right\}\mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right)[\boldsymbol{\Sigma}_z]_{(:,s)} = 0\end{aligned} \tag{E.67}$$

Note that the FOC holds for all sectors $s = 1, 2, \ldots N$, thus we can arrange equation (E.67) in a row vector for each sector s, and employ the notation in Lemma 4.3:

$$\left\{\boldsymbol{r}^o + \boldsymbol{r}^c + \boldsymbol{r}^d\right\}J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa}) + \left\{-\left(\frac{1}{(\gamma+1/\eta)} - 1\right)(\boldsymbol{\beta}'\mathbf{Q1})\,\boldsymbol{\beta}' - \boldsymbol{\lambda}'\Big(\operatorname{diag}(\mathbf{Q1}) - \operatorname{diag}(\mathbf{AQ1})\,\mathbf{A}\Big)\right\}\mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{\Sigma}_z = \mathbf{0}' \tag{E.68}$$

Now recall $J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})$ from Proposition 4.2,

$$J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa}) = \left[\mathbf{I} - \mathcal{T}_{\boldsymbol{\mu}^v}\right]^{-1}\mathcal{T}_{\boldsymbol{\kappa}} = \mathcal{R}\operatorname{diag}\left(\left\{\frac{2(1-\mu_i)\rho_i}{\mu_i V_i}\right\}_{i=1}^N\right)\phi\Sigma_{\mathbf{z}}. \tag{E.69}$$

Since $\mathbf{Q} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu})$, we obtain

$$\phi = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}(-\mathbf{I} + \boldsymbol{\alpha\kappa}) = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\left((\mathbf{I} - \boldsymbol{\mu}\mathbf{A}) - (\mathbf{I} - \boldsymbol{\mu})\right)\mathbf{L}(-\mathbf{I} + \boldsymbol{\alpha\kappa}) = (\mathbf{I} - \mathbf{Q})\mathbf{L}(-\mathbf{I} + \boldsymbol{\alpha\kappa}). \tag{E.70}$$

(E.70) serves as an important alternative representation of the pricing function. Using (E.69) and (E.70), equation (E.68) then simplifies to

$$\left[\left\{\boldsymbol{r}^o + \boldsymbol{r}^c + \boldsymbol{r}^d\right\}\mathcal{R}\operatorname{diag}\left(\left\{\frac{2(1-\mu_i)\rho_i}{\mu_i V_i}\right\}_{i=1}^N\right)(\mathbf{I} - \mathbf{Q}) + \left\{\left(\frac{1}{(\gamma+1/\eta)} - 1\right)(\boldsymbol{\beta}'\mathbf{Q1})\,\boldsymbol{\beta}' + \boldsymbol{\lambda}'\Big(\operatorname{diag}(\mathbf{Q1}) - \operatorname{diag}(\mathbf{AQ1})\,\mathbf{A}\Big)\right\}\mathbf{Q}\right]\mathbf{L}(-\mathbf{I} + \boldsymbol{\alpha\kappa})\boldsymbol{\Sigma}_z = \mathbf{0}', \tag{E.71}$$

where we apply matrix identities $\mathbf{L}\boldsymbol{\alpha} = \mathbf{1}$ and $(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}) = \mathbf{L}(\mathbf{I} - \boldsymbol{\alpha\kappa})$.

Note that in (E.71), $\boldsymbol{\Sigma}_z$ is invertible with diagonal terms $\sigma_i^2 > 0, \forall i = 1, 2, ...N$, thus we must have

$$\underbrace{\left[\left\{\boldsymbol{r}^o + \boldsymbol{r}^c + \boldsymbol{r}^d\right\}\mathcal{R}\,\mathrm{diag}\left(\left\{\frac{2(1-\mu_i)\rho_i}{\mu_i V_i}\right\}_{i=1}^N\right)(\mathbf{I}-\mathbf{Q}) + \left\{\left(\frac{1}{(\gamma+1/\eta)} - 1\right)(\boldsymbol{\beta}'\mathbf{Q1})\,\boldsymbol{\beta}' + \boldsymbol{\lambda}'\Big(\mathrm{diag}\,(\mathbf{Q1}) - \mathrm{diag}\,(\mathbf{AQ1})\,\mathbf{A}\Big)\right\}\mathbf{Q}\right]\mathbf{L}}_{\equiv\boldsymbol{\Phi}}(-\mathbf{I}+\boldsymbol{\alpha\kappa}) = \mathbf{0}', \tag{E.72}$$

where we define

$$\boldsymbol{\Phi} = \left\{\left(\frac{1}{(\gamma+1/\eta)} - 1\right)(\boldsymbol{\beta}'\mathbf{Q1})\,\boldsymbol{\beta}' + \boldsymbol{\lambda}'\Big(\mathrm{diag}\,(\mathbf{Q1}) - \mathrm{diag}\,(\mathbf{AQ1})\,\mathbf{A}\Big)\right\}\mathbf{QL} + \left\{\boldsymbol{r}^o + \boldsymbol{r}^c + \boldsymbol{r}^d\right\}\mathcal{R}\,\mathrm{diag}\left(\left\{\frac{2(1-\mu_i)\rho_i}{\mu_i V_i}\right\}_{i=1}^N\right)(\mathbf{I}-\mathbf{Q})\mathbf{L}$$

Therefore, write the above FOC as

$$\boldsymbol{\Phi} - \boldsymbol{\Phi\alpha\kappa} = \mathbf{0}'$$

Therefore, the optimal monetary policy rule is given by the fixed-point

$$\boldsymbol{\kappa} = \frac{\boldsymbol{\Phi}(\boldsymbol{\kappa})}{\boldsymbol{\Phi}(\boldsymbol{\kappa})\,\boldsymbol{\alpha}} \tag{E.73}$$

as desired. □

Step 2: Transformation to Price-Stabilization

We begin the proof of the second step by a lemma, which provides necessary and sufficient conditions on the existence of price-stabilization policy.

Lemma E.3. *The monetary policy rule under Proposition 3.6, which is supported by a nominal wage rule $w_t = \sum_{i=1}^N \kappa_i z_{it}$, can be implemented by a price-stabilization policy of the form*

$$\sum_{i=1}^N \varphi_i p_{it} = 0; \qquad \boldsymbol{\varphi} = \left(\varphi_1, \ldots, \varphi_N\right) \in \mathbb{R}^N,$$

if and only if row vector $\boldsymbol{\kappa}$ satisfies $\boldsymbol{\kappa\alpha} = 1$.

Proof. Note the existence of price stabilization is equivalent to a linear matrix equation,

$$\boldsymbol{\varphi\phi} = \mathbf{0}$$

which holds for all realizations of shocks $\boldsymbol{z}_t$. Simple linear algebra then implies such stabilization exists if and only if the row and columns of $\boldsymbol{\phi}$ are linear dependent ($\boldsymbol{\phi}$ is non-invertible). By Proposition 3.4,

$$\boldsymbol{\phi} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\,\boldsymbol{\mu}\,(-\mathbf{I} + \boldsymbol{\alpha\kappa})$$

is non-invertible in the interior solution of $\boldsymbol{\mu}$ if and only if

$$\det(-\mathbf{I} + \boldsymbol{\alpha\kappa}) = 0. \tag{E.74}$$

Now apply the matrix determinant lemma,

$$\det(-\mathbf{I} + \boldsymbol{\alpha\kappa}) = (1 - \boldsymbol{\kappa\alpha})\det(-\mathbf{I}) = -(1 - \boldsymbol{\kappa\alpha}).$$

Therefore, a price-stabilization policy exists if and only if $\boldsymbol{\kappa\alpha} = 1$, as desired. □

Obviously, the optimal montery policy rule in equation (E.65) and (E.73) satisfies the condition $\boldsymbol{\kappa\alpha} = 1$, thus the optimal monetary policy can be achieved through a price stabilization policy. By equations (E.55) and (E.70), we notice that two matrix identities hold:

$$\boldsymbol{\phi} = (\mathbf{I} - \mathbf{Q})\mathbf{L}(-\mathbf{I} + \boldsymbol{\alpha\kappa}); \qquad \mathbf{Q}\left(\mathbf{1}\boldsymbol{\kappa} - \mathbf{L}\right) = \mathbf{L}\left(\boldsymbol{\mu}^{-1} - \mathbf{I}\right)\boldsymbol{\phi},$$

where we use the relation $(\mathbf{I} - \mathbf{A})\mathbf{1} = \boldsymbol{\alpha}$. Using these two identities in (E.72), we obtain

$$\left[\left\{\boldsymbol{r}^o + \boldsymbol{r}^c + \boldsymbol{r}^d\right\}\mathcal{R}\,\mathrm{diag}\left(\left\{\frac{2(1-\mu_i)\rho_i}{\mu_i V_i}\right\}_{i=1}^N\right) + \left\{\left(\frac{1}{(\gamma + 1/\eta)} - 1\right)(\boldsymbol{\beta}'\mathbf{Q}\mathbf{1})\,\boldsymbol{\beta}' + \boldsymbol{\lambda}'\Big(\mathrm{diag}\,(\mathbf{Q}\mathbf{1}) - \mathrm{diag}\,(\mathbf{A}\mathbf{Q}\mathbf{1})\,\mathbf{A}\Big)\right\}\mathbf{L}\left(\boldsymbol{\mu}^{-1} - \mathbf{I}\right)\right]\boldsymbol{\phi} = \mathbf{0}' \tag{E.75}$$

Therefore, in the spirit of Lemma E.3, the price-stabilization weight is given by

$$\boldsymbol{\varphi} = \left\{\left(\frac{1}{(\gamma + 1/\eta)} - 1\right)(\boldsymbol{\beta}'\mathbf{Q}\mathbf{1})\,\boldsymbol{\beta}' + \boldsymbol{\lambda}'\Big(\mathrm{diag}\,(\mathbf{Q}\mathbf{1}) - \mathrm{diag}\,(\mathbf{A}\mathbf{Q}\mathbf{1})\,\mathbf{A}\Big)\right\}\mathbf{L}(\boldsymbol{\mu}^{-1} - \mathbf{I}) + \left\{\boldsymbol{r}^o + \boldsymbol{r}^c + \boldsymbol{r}^d\right\}\mathcal{R}\,\mathrm{diag}\left(\left\{\frac{2\rho_i}{V_i}\right\}_{i=1}^N\right)(\boldsymbol{\mu}^{-1} - \mathbf{I}).$$

Recall from Proposition E.2 that we define the matrix of general equilibrium propagation (feedbacks) in attentions as

$$\mathcal{R} \equiv \left[I - \mathcal{T}_{\mu^v}\right]^{-1}.$$

Next, we decompose the price-stabilization policy vector $\boldsymbol{\varphi}$ into two components: $\boldsymbol{\varphi}^e = (\varphi_1^e, \varphi_2^e, ...\varphi_N^e)$ represents policy response due to endogenous changes in attention (price flexibilities), and $\boldsymbol{\varphi}^x = (\varphi_1^x, \varphi_2^x, ...\varphi_N^x)$ represents policy response holding the nominal rigidities constant; that is

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}^e + \boldsymbol{\varphi}^x$$

The two components then follows

$$\boldsymbol{\varphi}^e = \left\{\boldsymbol{r}^o + \boldsymbol{r}^c + \boldsymbol{r}^d\right\}\mathcal{R}\,\mathrm{diag}\left(\left\{\frac{2\rho_i}{V_i}\right\}_{i=1}^N\right)(\boldsymbol{\mu}^{-1} - \mathbf{I}); \tag{E.76}$$

$$\varphi^x = \left\{ \left(\frac{1}{(\gamma + 1/\eta)} - 1 \right) (\boldsymbol{\beta}'\mathbf{Q1})\, \boldsymbol{\beta}' + \boldsymbol{\lambda}' \Big(\operatorname{diag}(\mathbf{Q1}) - \operatorname{diag}(\mathbf{AQ1})\, \mathbf{A} \Big) \right\} \mathbf{L}(\boldsymbol{\mu}^{-1} - \mathbf{I}). \tag{E.77}$$

Step 3: Algebraic Details to Scalar Representation (4.4) and (4.13)

First, we write the endogenous components (E.76) in scalar form as

$$\varphi_i^e = 2 \left(\sum_{j=1}^{N} \left[r_j^o + r_j^c + r_j^d \right] r_{ji} \right) \left(\frac{1}{\mu_i} - 1 \right) \frac{\rho_i}{\mathbb{V}(\mathrm{mc}_{it})}, \tag{E.78}$$

where r_j^o, r_j^c, r_j^d, and r_{ji} are elements of $\boldsymbol{r}^o$, $\boldsymbol{r}^c$, $\boldsymbol{r}^d$, and $\mathcal{R}$, respectively. Note that if we employ the notation in (E.62) and (E.67), we can express the weight in more details,

$$\varphi_i^e = 2 \left(\sum_{j=1}^{N} \left[\frac{1}{2}\chi_j + \frac{1}{(\gamma + 1/\eta)} r_j^{\beta'} + \sum_{k=1}^{N} \lambda_k \left(r_{kj}^I - r_{kj}^A \right) - r_j^{\beta'} \right] r_{ji} \right) \left(\frac{1}{\mu_i} - 1 \right) \frac{\rho_i}{\mathbb{V}(\mathrm{mc}_{it})}$$

where r_{ji}, r_{kj}^I, r_{kj}^A and $r_j^{\beta'}$ are elements of $\mathcal{R}$, $\mathbf{R}^{\mathbf{I}}$, $\mathbf{R}^{\mathbf{A}}$ and $\mathbf{r}^{\beta'}$, respectively.

Next, we work on the exogenous component (E.77). By definition, $\boldsymbol{\rho} = (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}\boldsymbol{\alpha}$ and $\mathbf{Q} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu})$. Consequently,

$$\mathbf{Q1} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}(\mathbf{I} - \boldsymbol{\mu})\mathbf{L}\boldsymbol{\alpha} = \left(\mathbf{I} - (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}(\mathbf{I} - \mathbf{A})\right) \mathbf{L}\boldsymbol{\alpha} = \left(\mathbf{L} - \boldsymbol{\mu}(\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}\right) \boldsymbol{\alpha} = \mathbf{1} - \boldsymbol{\mu}\boldsymbol{\rho} \tag{E.79}$$

where I use the matrix identities $\mathbf{L}\boldsymbol{\alpha} = \mathbf{1}$ and (D.36). We define $\rho_0 = \boldsymbol{\beta}'\boldsymbol{\mu}\boldsymbol{\rho}$, which implies that $\boldsymbol{\beta}'\mathbf{Q1} = 1 - \rho_0$ as $\sum_{i=1}^N \beta_i = 1$.

Therefore, the component of the policy response holding the nominal rigidities constant can be further simplified as

$$\varphi^x = \boldsymbol{\lambda}' \left\{ \left(\frac{1}{(\gamma + 1/\eta)} - 1 \right) (1 - \rho_0)\, \mathbf{I} + \Big(\mathbf{I} - \boldsymbol{\mu} \operatorname{diag}(\boldsymbol{\rho}) - \operatorname{diag}(\mathbf{A1} - \mathbf{A}\boldsymbol{\mu}\boldsymbol{\rho})\, \mathbf{A} \Big) \mathbf{L} \right\} (\boldsymbol{\mu}^{-1} - \mathbf{I}),$$

where I use $\boldsymbol{\beta}'\mathbf{L} = \boldsymbol{\lambda}'$ from (D.5). Next, we derive a matrix identity

$$\mathbf{A}\boldsymbol{\mu}\boldsymbol{\rho} = (\mathbf{I} - (\mathbf{I} - \mathbf{A}\boldsymbol{\mu}))\, (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}\boldsymbol{\alpha} = (\mathbf{I} - \mathbf{A}\boldsymbol{\mu})^{-1}\boldsymbol{\alpha} - \boldsymbol{\alpha} = \boldsymbol{\rho} - (\mathbf{I} - \mathbf{A})\mathbf{1}$$

where $\boldsymbol{\alpha} = (\mathbf{I} - \mathbf{A})\mathbf{1}$. We then obtain

$$\begin{aligned}
\Big(\mathbf{I} - \boldsymbol{\mu} \operatorname{diag}(\boldsymbol{\rho}) - \operatorname{diag}(\mathbf{A1} - \mathbf{A}\boldsymbol{\mu}\boldsymbol{\rho})\, \mathbf{A} \Big) \mathbf{L} &= \Big(\mathbf{I} - \boldsymbol{\mu} \operatorname{diag}(\boldsymbol{\rho}) - \mathbf{A} + \operatorname{diag}(\boldsymbol{\rho})\, \mathbf{A} \Big) \mathbf{L} \\
&= \Big((\mathbf{I} - \mathbf{A}) + (\mathbf{I} - \boldsymbol{\mu}) \operatorname{diag}(\boldsymbol{\rho}) - \operatorname{diag}(\boldsymbol{\rho})\, (\mathbf{I} - \mathbf{A}) \Big) \mathbf{L} \\
&= \mathbf{I} + (\mathbf{I} - \boldsymbol{\mu}) \operatorname{diag}(\boldsymbol{\rho})\, \mathbf{L} - \operatorname{diag}(\boldsymbol{\rho})
\end{aligned}$$

Consequently, the component of the policy response holding the nominal rigidities constant simplifies to

$$\boldsymbol{\varphi}^x = \boldsymbol{\lambda}' \left\{ \left(\frac{(1-\rho_0)}{(\gamma + 1/\eta)} + \rho_0 \right) \mathbf{I} + (\mathbf{I} - \boldsymbol{\mu}) \operatorname{diag}(\boldsymbol{\rho}) \mathbf{L} - \operatorname{diag}(\boldsymbol{\rho}) \right\} (\boldsymbol{\mu}^{-1} - \mathbf{I}) \tag{E.80}$$

It follows the exogenous component of weight in scalar form is given by

$$\varphi_i^x = \left[\frac{(1-\rho_0)}{(\gamma + 1/\eta)} \lambda_i + \sum_{j=1}^{N} (1-\mu_i) \lambda_j \rho_j l_{ji} + (\rho_0 - \rho_i) \lambda_i \right] \left(\frac{1}{\mu_i} - 1 \right) \tag{E.81}$$

Finally, we modify the formula (E.78) and (E.81) by adding and subtracting the exogenous component associated with the within-sector price dispersion under the exogenous-information (fixed-capacity) model, that is, the last component in (4.4): $\mu_i \lambda_i \theta_i \rho_i \left(\frac{1}{\mu_i} - 1 \right)$. In this case, the exogenous component exactly align with its counterpart under exogenous-information:

$$\varphi_i^x = \left[\frac{(1-\rho_0)}{(\gamma + 1/\eta)} \lambda_i + \sum_{j=1}^{N} (1-\mu_i) \lambda_j \rho_j l_{ji} + (\rho_0 - \rho_i) \lambda_i + \mu_i \lambda_i \theta_i \rho_i \right] \left(\frac{1}{\mu_i} - 1 \right) \tag{E.82}$$

Now (E.82) and (4.4) coincide. The modified endogenous component then follows,

$$\begin{aligned}
\varphi_i^e &= 2 \left(\sum_{j=1}^{N} \left[r_j^o + r_j^c + r_j^d \right] r_{ji} \right) \left(\frac{1}{\mu_i} - 1 \right) \frac{\rho_i}{\mathbb{V}(\mathrm{mc}_{it})} - \mu_i \lambda_i \theta_i \rho_i \left(\frac{1}{\mu_i} - 1 \right) \\
&= \left\{ 2 \left(\sum_{j=1}^{N} \left[r_j^o + r_j^c + r_j^d \right] r_{ji} \right) - \mu_i \lambda_i \theta_i \mathbb{V}(\mathrm{mc}_{it}) \right\} \left(\frac{1}{\mu_i} - 1 \right) \frac{\rho_i}{\mathbb{V}(\mathrm{mc}_{it})} \\
&= \left\{ 2 \left(\sum_{j=1}^{N} \left[r_j^o + r_j^c + r_j^d \right] r_{ji} \right) - (\lambda_i \theta_i \mathbb{V}(\mathrm{mc}_{it}) - \chi_i) \right\} \frac{\rho_i}{\mu_i} \frac{1-\mu_i}{\mathbb{V}(\mathrm{mc}_{it})}
\end{aligned}$$

where the last equality follows from the equilibrium fixed-point (3.6). Hence, we write the endogenous policy weight as

$$\varphi_i^e = \left\{ 2 \left(\sum_{j=1}^{N} \left[r_j^o + r_j^c + r_j^d \right] r_{ji} \right) - (\lambda_i \theta_i \mathbb{V}(\mathrm{mc}_{it}) - \chi_i) \right\} \frac{\rho_i}{\mu_i} \frac{1-\mu_i}{\mathbb{V}(\mathrm{mc}_{it})} \tag{E.83}$$

as desired. The proof is now complete. □

E.7 Proof of Proposition 4.4

Proof. Recall that the output-gap volatility in terms of cross-sectional average pricing errors is given by (E.2). We first state the optimal output-gap volatility optimization problem in the following definition.

Definition 5. *The central bank designs optimal OG policy by solving the following constrained optimization problem*

$$\min_{\{\boldsymbol{\mu}^v,\boldsymbol{\kappa}\}} \mathbb{V}\left(c_t - c_t^*\right) = \mathbb{E}\left[\frac{1}{(\gamma+1/\eta)^2}\left(\sum_{j=1}^{n}\beta_j \bar{e}_{jt}\right)^2\right]$$

subject to the equilibrium fixed-point for each sector $i = 1, 2...N$*,*

$$\mu_i = \max\left\{0, 1 - \frac{\chi_i}{\theta_i\lambda_i\mathbb{V}(mc_{it})}\right\}; \quad \mathbb{V}(mc_{it}) = \left\|\mathbf{e}_i\left[(\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1}(-\mathbf{I}+\boldsymbol{\alpha\kappa})\right]\boldsymbol{\Sigma}_z^{1/2}\right\|^2.$$

In order to obtain the optimal OG policy, we need to solve the first-order condition of the expected output-gap volatility on the monetary policy rule,

$$\frac{d\mathbb{E}\left[\Sigma_t\right]}{d\kappa_s} = \frac{1}{(\gamma+1/\eta)^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\beta_i\beta_j\mathbb{E}\left[\bar{e}_{it}\frac{d\bar{e}_{jt}}{d\kappa_s}\right] = 0; \Leftrightarrow \mathbf{r}^{\beta'}\left[J_{\boldsymbol{\mu}^v}(\boldsymbol{\kappa})\right]_{(:,s)} - (\boldsymbol{\beta}'\mathbf{Q}\mathbf{1})\,\boldsymbol{\beta}'\mathbf{Q}\,(\mathbf{L}-\mathbf{1}\boldsymbol{\kappa})\,[\boldsymbol{\Sigma}_z]_{(:,s)} = 0$$

where the second identity follows from (E.57), and $\mathbf{r}^{\beta'} = [(\boldsymbol{\beta}'\mathbf{L}\,(\boldsymbol{\mu}-\mathbf{I})\,\mathbb{V}(\mathrm{mc}_{it}))\odot(\boldsymbol{\beta}'\mathbf{H})]$. Applying the same methods we employed to derive (E.72), the above FOC condition can be simplified to

$$\left[\mathbf{r}^{\beta'}\mathcal{R}\,\mathrm{diag}\left(\left\{\frac{2(1-\mu_i)\rho_i}{\mu_i V_i}\right\}_{i=1}^{N}\right)(\mathbf{I}-\mathbf{Q}) + (\boldsymbol{\beta}'\mathbf{Q}\mathbf{1})\,\boldsymbol{\beta}'\mathbf{Q}\right]\mathbf{L}(-\mathbf{I}+\boldsymbol{\alpha\kappa})\boldsymbol{\Sigma}_z = \mathbf{0}'. \tag{E.84}$$

Therefore, the optimal OG policy rule is given by the fixed-point

$$\boldsymbol{\kappa}^{OG} = \frac{\boldsymbol{\Phi}^{OG}}{\boldsymbol{\Phi}^{OG}\boldsymbol{\alpha}}$$

where we define

$$\boldsymbol{\Phi}^{OG} = \mathbf{r}^{\beta'}\mathcal{R}\,\mathrm{diag}\left(\left\{\frac{2(1-\mu_i)\rho_i}{\mu_i V_i}\right\}_{i=1}^{N}\right)(\mathbf{I}-\mathbf{Q})\mathbf{L} + (\boldsymbol{\beta}'\mathbf{Q}\mathbf{1})\,\boldsymbol{\beta}'\mathbf{Q}\mathbf{L}$$

Next, recall from Lemma E.3, the optimal OG policy can be implemented through a price-stabilization policy as

$$\begin{aligned}\boldsymbol{\varphi}^{OG} &= \left[\mathbf{r}^{\beta'}\mathcal{R}\,\mathrm{diag}\left(\left\{\frac{2(1-\mu_i)\rho_i}{\mu_i V_i}\right\}_{i=1}^{N}\right)(\mathbf{I}-\mathbf{Q})\mathbf{L} + (\boldsymbol{\beta}'\mathbf{Q}\mathbf{1})\,\boldsymbol{\beta}'\mathbf{Q}\mathbf{L}\right]\boldsymbol{\mu}^{-1}(\mathbf{I}-\boldsymbol{\mu}\mathbf{A}) \\ &= \underbrace{\mathbf{r}^{\beta'}\mathcal{R}\,\mathrm{diag}\left(\left\{\frac{2(1-\mu_i)\rho_i}{\mu_i V_i}\right\}_{i=1}^{N}\right)}_{\boldsymbol{\varphi}^{e,OG}} + \underbrace{(1-\rho_0)\,\boldsymbol{\lambda}'\left(\boldsymbol{\mu}^{-1}-\mathbf{I}\right)}_{\boldsymbol{\varphi}^{x,OG}}\end{aligned} \tag{E.85}$$

where the last identity follows from matrix identities[45] $\mathbf{QL} = \mathbf{L}\left(\boldsymbol{\mu}^{-1} - \mathbf{I}\right)(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}$, $(\mathbf{I} - \mathbf{Q})\mathbf{L} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}$, $\boldsymbol{\beta}'\mathbf{Q1} = 1 - \rho_0$, and $\boldsymbol{\beta}'\mathbf{L} = \boldsymbol{\lambda}'$. Equation (E.85) shows the price-stabilization policy can be splited into two components under the OG optimization problem, which is similar to our approach in Proposition 4.3. The next proposition characterizes a condition under which the optimal OG policy satisfies $\boldsymbol{\varphi}^{e,OG} \equiv \mathbf{0}$.

Proposition E.3. *If the interior optimal OG policy completely stabilizes output-gap volatility ($\mathbb{V}\left(c_t - c_t^*\right) = 0$), at the optimum the price-stabilization policy in (E.85) does not respond to endogenous changes in attention (price flexibilities); that is, $\boldsymbol{\varphi}^{e,OG} \equiv \mathbf{0}$.*

Proof. In matrix form, the expected output-gap volatility can be expressed as

$$\mathbb{V}\left(c_t - c_t^*\right) = \frac{1}{(\gamma + 1/\eta)^2}\boldsymbol{\beta}'\boldsymbol{\Sigma}_e\boldsymbol{\beta} = \left\|\boldsymbol{\beta}'\mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{\Sigma}_z^{1/2}\right\|^2 \geq 0$$

where $\boldsymbol{\Sigma}_e = \mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{\Sigma}_z\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right)'\mathbf{Q}'$ is defined as the covariance matrix of the cross-sectional average of sectoral pricing errors. Under the interior optimal OG policy, if the output-gap volatility vanishes, then $\left(\boldsymbol{\beta}'\mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{\Sigma}_z^{1/2}\right) = \mathbf{0}'$. Then at the optimum, since $\boldsymbol{\Sigma}_z^{1/2} \succ 0$ is invertible,

$$\mathbf{r}^{\beta'} = \left[(\boldsymbol{\beta}'\mathbf{L}\left(\boldsymbol{\mu} - \mathbf{I}\right)\mathbb{V}(\mathrm{mc}_{it})) \odot (\boldsymbol{\beta}'\mathbf{H})\right] = \left[\left(\left(\boldsymbol{\beta}'\mathbf{Q}\left(\mathbf{L} - \mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{\Sigma}_z^{1/2}\right)\boldsymbol{\Sigma}_z^{1/2}\boldsymbol{\phi}'\boldsymbol{\mu}^{-1}\right) \odot (\boldsymbol{\beta}'\mathbf{H})\right] \equiv \mathbf{0}$$

by matrix identity (E.56). Therefore, by (E.84) the optimal OG policy rule reduces to

$$\boldsymbol{\kappa}^{OG} = \frac{\boldsymbol{\beta}'\mathbf{QL}}{\boldsymbol{\beta}'\mathbf{Q1}}$$

As a result, $\boldsymbol{\varphi}^{e,OG} \equiv \mathbf{0}$ and the endogenous part of the price-stabilizaiton condition in (E.85) vanishes. □

Given the above property of the optimal OG policy, the price-stabilization now simplifies to

$$\boldsymbol{\varphi}^{OG} = \boldsymbol{\varphi}^{x,OG} = \left(1 - \rho_0\right)\boldsymbol{\lambda}'\left(\boldsymbol{\mu}^{-1} - \mathbf{I}\right); \qquad \boldsymbol{\varphi}^{e,OG} \equiv \mathbf{0},$$

which can be expressed as scalar form,

$$\varphi_i = \varphi_i^{x,OG} = \frac{(1 - \rho_0)}{(\gamma + 1/\eta)}\lambda_i\left(\frac{1}{\mu_i} - 1\right); \qquad \varphi_i^{e,OG} \equiv 0; \qquad i = 1, 2, ...N.$$

The proof is now complete. □

[45] Specifically, we notice that by (E.55) $\mathbf{QL} = \mathbf{L}(\boldsymbol{\mu}^{-1} - \mathbf{I})(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}$. Combining this equation with (D.36) yields

$$(\mathbf{I} - \mathbf{Q})\mathbf{L} = \mathbf{L} - \mathbf{L}(\boldsymbol{\mu}^{-1} - \mathbf{I})(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu} = \mathbf{L}\left(\boldsymbol{\mu}^{-1}(\mathbf{I} - \boldsymbol{\mu}\mathbf{A}) - (\boldsymbol{\mu}^{-1} - \mathbf{I})\right)(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu} = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}.$$

The matrix identities $\boldsymbol{\beta}'\mathbf{Q1} = 1 - \rho_0$ follows from (E.79) and $\boldsymbol{\beta}'\mathbf{L} = \boldsymbol{\lambda}'$ from (D.5).

E.8 Proof of Proposition 4.5

Proof. Under rational inattention, when only minimizing welfare loss associated with the within-sector price dispersion. The objective function is simply

$$L^{within} = \frac{1}{2}\boldsymbol{\chi}'\boldsymbol{\mu}^{v}$$

Therefore, optimal monetary policy only operate through the endogenous-attention channel. The mathematics is identical to our previous derivations, and the optimal policy weight is simply

$$\varphi_i^{within} = \left(\sum_{j=1}^{N} \chi_j \mathcal{R}_{ji}\right) \frac{\rho_i}{\mathbb{V}(\mathrm{mc}_{it})} \left(\frac{1}{\mu_i} - 1\right). \tag{E.86}$$

The proof is complete.

□

F. Additional Theory and Policy Results

F.1 Equivalence between Attention and Signal Precision (Sectoral Nominal Rigidity)

In this subsection, we establish equivalence between signal precision of sector i, μ_i, and the total attention capacity chosen by sector i. Let δ_i^* be the endogenous amount of information capacity chosen by firms, measured in terms of nats. By definition,

$$\delta_i^* = \frac{1}{2}\log\det\mathbf{\Sigma}_z - \frac{1}{2}\log\det\mathbf{\Sigma}_{z|x_i}$$

The following proposition establishes a monotone, one-to-one mapping between δ_i^* and μ_i.

Proposition F.1. *The endogenous attention (total entropy reduction) δ_i^* chosen by firms in sector i is monotonically increasing with μ_i and is given by:*

$$\delta_i^* = -\frac{1}{2}\log\left(1-\mu_i\right).$$

Proof. Using our previous result (D.25) and partition $\mathbf{U}_i = \begin{bmatrix}\zeta_1^i & \zeta_2^i\end{bmatrix}$ where eigenvector ζ_1^i corresponds to d_1^i defined in Appendix D.3,

$$\begin{aligned}
\mathbf{\Sigma}_{z|x_i} &= \mathbf{\Sigma}_z^{\frac{1}{2}}\begin{bmatrix}\zeta_1^i & \zeta_2^i\end{bmatrix}\begin{bmatrix}\frac{\chi_i}{2d_1^i} & 0\\ & \\ 0 & \mathbf{I}_{N-1}\end{bmatrix}\begin{bmatrix}\zeta_1^i & \zeta_2^i\end{bmatrix}'\mathbf{\Sigma}_z^{\frac{1}{2}}\\
&= \mathbf{\Sigma}_z^{\frac{1}{2}}\left[\zeta_1^i\zeta_1^{i,T} + \zeta_2^i\zeta_2^{i,T} - \left(1-\frac{\chi_i}{2d_1^i}\right)\zeta_1^i\zeta_1^{i,T}\right]\mathbf{\Sigma}_z^{\frac{1}{2}}\\
&= \mathbf{\Sigma}_z^{\frac{1}{2}}\left[\mathbf{I} - \left(1-\frac{\chi_i}{2d_1^i}\right)\frac{\mathbf{\Sigma}_z^{\frac{1}{2}}\mathbf{\Omega}_{iz}\mathbf{\Sigma}_z^{\frac{1}{2}}}{d_1^i}\right]\mathbf{\Sigma}_z^{\frac{1}{2}}\\
&= \mathbf{\Sigma}_z - \left[\frac{\left(2d_1^i-\chi_i\right)}{2(d_1^i)^2}\right]\left(\mathbf{\Sigma}_z\mathbf{\Omega}_{iz}\mathbf{\Sigma}_z\right)\\
&= \mathbf{\Sigma}_z - \frac{\chi_i}{2R_i d_1^i \nu_i^2}\left(\mathbf{\Sigma}_z\mathbf{\Omega}_{iz}\mathbf{\Sigma}_z\right)
\end{aligned}$$

where we use the definition of unitary matrix $\zeta_1^i\zeta_1^i)' + \zeta_2^i(\zeta_2^i)' = \mathbf{U}_i\mathbf{U}_i' = \mathbf{I}$ and the derivation of endogenous noise variance in Appendix D.3. Next, we compute the ratio of determinants of prior and posterior covariances,

$$\frac{\det\mathbf{\Sigma}_{z|x_i}}{\det\mathbf{\Sigma}_z} = \det\mathbf{\Sigma}_{z|x_i}\mathbf{\Sigma}_z^{-1} = \det\left(\mathbf{I} - \frac{\chi_i}{2R_i d_1^i\nu_i^2}\mathbf{\Sigma}_z\mathbf{\Omega}_{iz}\right) = \det\left(\mathbf{I} - \frac{\chi_i}{2R_i d_1^i\nu_i^2}\mathbf{\Sigma}_z\mathbf{G}_i'\mathbf{G}_i\right),$$

since $\mathbf{\Sigma}_z\mathbf{G}_i'$ is a $N\times 1$ vector while $\mathbf{G}_i$ is a $1\times N$ vector, we employ the matrix determinant lemma for outer product

$\mathbf{\Sigma}_z \mathbf{G}_i' \mathbf{G}_i$,

$$\det\left(\mathbf{I} - \frac{\chi_i}{2R_i d_1^i \nu_i^2} \mathbf{\Sigma}_z \mathbf{G}_i' \mathbf{G}_i\right) = 1 - \frac{\chi_i}{2R_i d_1^i \nu_i^2} \mathbf{G}_i \mathbf{\Sigma}_z \mathbf{G}_i' = 1 - \frac{\chi_i}{2d_1^i \nu_i^2} \mathbb{V}\left(\mathrm{mc}_{it}\right) = 1 - \frac{\mathbb{V}\left(\mathrm{mc}_{it}\right)}{\mathbb{V}\left(\mathrm{mc}_{it}\right) + \nu_i^2} = 1 - \mu_i$$

where the second equality follows from $\mathbb{V}\left(\mathrm{mc}_{it}\right) = R_i^{-1} \parallel \mathbf{\Sigma}_z^{\frac{1}{2}} \mathbf{G}_i' \parallel^2$ in (D.19). The last two equalities follows from previous derivations of d_1^i, ν_i^2, and μ_i: (D.20), (D.27), and (D.28). Therefore, it follows that

$$\delta_i^* = -\frac{1}{2} \log\left(1 - \mu_i\right)$$

which completes the proof. □

To summarize, the isomorphic relation between δ_i^* and μ_i indicates that sectoral attentions and signal precisions, and therefore nominal rigidities, are indeed equivalent.

F.2 An Illustrative Example

To further clarify the dependence of sectoral attentions and endogenous feedbacks of attentions on model primitives $(\mathbf{A}, \mathbf{\Sigma}_{\boldsymbol{z}}, \boldsymbol{\kappa})$, we present a special example with closed-form characterizations. In this special example, we isolate a particular channel of endogenous feedbacks of attentions documented in Proposition 4.2 , which corresponds to the within-sector streategic complementarity, or the diagonal elements in $\mathcal{T}_\mu$ (see Corollary F.1). Specifically, we consider an economy in which a sector i is rationally inattentive while other sectors have complete information. That is, we impose the following assumption on sectoral information cost,

Assumption 2. *$\chi_i > 0$ and $\chi_j = 0, \forall j \neq i$.*

Assumption 2 implies that the rest of economies feature full price-flexibilities $\mu_j = 1, \forall j \neq i$. The convinent feature allows for closed-form characterization in the next proposition.

Proposition F.2. *Under Assumption 2, sector i's equilibrium attention μ_i is determined by the following quadratic equation*

$$\frac{\chi_i}{\theta_i \lambda_i} \frac{\left(1 + \left(1 - \mu_i\right)\left[\sum_{j=1}^N \left(l_{ij} a_{ji}\right)\right]\right)^2}{\sum_{j=1}^N \left[\left(\kappa_j - l_{ij}\right) \sigma_j\right]^2} = 1 - \mu_i$$

where $l_{ij} \geq 0$ is the (i, j) th element of the Leontief inverse matrix $\mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1}$. If $\sum_{j=1}^N l_{ij} a_{ji} < 1$ and $\frac{\theta_i \lambda_i}{\chi_i} \sum_{j=1}^N \left[\left(\kappa_j - l_{ij}\right) \sigma_j\right]^2 > \left[1 + \sum_{j=1}^N \left(l_{ij} a_{ji}\right)\right]^2$, the unique equilibrium solution is given by

$$\mu_i = 1 - \frac{\frac{\theta_i \lambda_i}{\chi_i} \sum_{j=1}^N \left[\left(\kappa_j - l_{ij}\right) \sigma_j\right]^2 - 2 \sum_{j=1}^N \left(l_{ij} a_{ji}\right) - \sqrt{\frac{\theta_i \lambda_i}{\chi_i} \sum_{j=1}^N \left[\left(\kappa_j - l_{ij}\right) \sigma_j\right]^2 \left[\frac{\theta_i \lambda_i}{\chi_i} \sum_{j=1}^N \left[\left(\kappa_j - l_{ij}\right) \sigma_j\right]^2 - 4 \sum_{j=1}^N (l_{ij} a_{ji})\right]}}{2\left(\sum_{j=1}^N \left(l_{ij} a_{ji}\right)\right)^2} \tag{F.1}$$

Proof. Under Assumption 2, the attention-distorted Leontief matrix can be expanded as

$$(\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1} = \left(\mathbf{I}-\mathbf{A}\begin{bmatrix}1 & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ \vdots & \dots & \mu_i & \vdots\\ 0 & \dots & \ddots & 1\end{bmatrix}\right)^{-1} = \left(\mathbf{I}-\mathbf{A}-\mathbf{A}\begin{bmatrix}0 & 0 & \dots & 0\\ 0 & 0 & \dots & 0\\ \vdots & \dots & (\mu_i-1) & \vdots\\ 0 & \dots & \ddots & 0\end{bmatrix}\right)^{-1} = \left(\mathbf{I}-\mathbf{A}-\mathbf{A}\mathbf{G}_{\mu_i}\right)^{-1},$$

where we define $\mathbf{G}_{\mu_i} = \begin{bmatrix}0 & 0 & \dots & 0\\ 0 & 0 & \dots & 0\\ \vdots & \dots & (\mu_i-1) & \vdots\\ 0 & \dots & \ddots & 0\end{bmatrix}$. Then by matrix geometric series expansion,

$$(\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1} = \left(\mathbf{I}-\mathbf{A}-\mathbf{A}\mathbf{G}_{\mu_i}\right)^{-1} = \sum_{n=0}^{\infty}\left(\mathbf{L}\mathbf{A}\mathbf{G}_{\mu_i}\right)^{n}\mathbf{L},$$

where $\mathbf{L} = (\mathbf{I}-\mathbf{A})^{-1}$ is the Leontief inverse. Expanding the summation,

$$\sum_{n=0}^{\infty}\left(\mathbf{L}\mathbf{A}\mathbf{G}_{\mu_i}\right)^{n} = \mathbf{I} + \begin{bmatrix}0 & 0 & \dots & \left[(\mu_i-1)(\mathbf{L}_1^r\mathbf{A}_i^c)\right]\sum_{n=0}^{\infty}\left[(\mu_i-1)(\mathbf{L}_i^r\mathbf{A}_i^c)\right]^n & 0 & \dots & 0\\ 0 & 0 & \dots & \left[(\mu_i-1)(\mathbf{L}_2^r\mathbf{A}_i^c)\right]\sum_{n=0}^{\infty}\left[(\mu_i-1)(\mathbf{L}_i^r\mathbf{A}_i^c)\right]^n & 0 & \dots & 0\\ \vdots & \dots & \dots & \left[(\mu_i-1)(\mathbf{L}_i^r\mathbf{A}_i^c)\right]\sum_{n=0}^{\infty}\left[(\mu_i-1)(\mathbf{L}_i^r\mathbf{A}_i^c)\right]^n & 0 & \dots & 0\\ \vdots & \dots & \dots & \left[(\mu_i-1)(\mathbf{L}_N^r\mathbf{A}_i^c)\right]\sum_{n=0}^{\infty}\left[(\mu_i-1)(\mathbf{L}_i^r\mathbf{A}_i^c)\right]^n & 0 & \dots & 0\end{bmatrix},$$

where $\mathbf{L}_n^r$ denotes the nth row of the Leontief matrix $\mathbf{L}$, and $\mathbf{A}_i^c$ denotes the ith column of $\mathbf{A}$. The inner product is defined as $\mathbf{L}_n^r\mathbf{A}_i^c = \sum_{j=1}^{N}\left(l_{nj}a_{ji}\right)$. Since the matrix inversion $(\mathbf{I}-\mathbf{A}\boldsymbol{\mu})^{-1}$ exists, the scalar infinite series converge to

$$\sum_{n=0}^{\infty}\left[(\mu_i-1)(\mathbf{L}_i^r\mathbf{A}_i^c)\right]^n = \frac{1}{\left(1-(\mu_i-1)\left[\sum_{j=1}^{N}\left(l_{ij}a_{ji}\right)\right]\right)}.$$

Therefore, the matrix series converges to $\sum_{n=0}^{\infty}\left(\mathbf{LAG}_{\mu_i}\right)^n = \begin{bmatrix} 1 & 0 & \dots & \frac{(\mu_i-1)\left[\sum_{j=1}^{N}\left(l_{1j}a_{ji}\right)\right]}{\left(1-(\mu_i-1)\left[\sum_{j=1}^{N}\left(l_{ij}a_{ji}\right)\right]\right)} & 0 & \dots & 0 \\ 0 & 1 & \dots & \frac{(\mu_i-1)\left[\sum_{j=1}^{N}\left(l_{2j}a_{ji}\right)\right]}{\left(1-(\mu_i-1)\left[\sum_{j=1}^{N}\left(l_{ij}a_{ji}\right)\right]\right)} & 0 & \dots & 0 \\ \vdots & \dots & \dots & 1+\frac{(\mu_i-1)\left[\sum_{j=1}^{N}\left(l_{ij}a_{ji}\right)\right]}{\left(1-(\mu_i-1)\left[\sum_{j=1}^{N}\left(l_{ij}a_{ji}\right)\right]\right)} & 0 & \dots & 0 \\ \vdots & \dots & \dots & \frac{(\mu_i-1)\left[\sum_{j=1}^{N}\left(l_{Nj}a_{ji}\right)\right]}{\left(1-(\mu_i-1)\left[\sum_{j=1}^{N}\left(l_{ij}a_{ji}\right)\right]\right)} & 0 & \dots & 1 \end{bmatrix}$.

Define the RI-induced distortion coefficient as

$$\varpi_i = 1 + \frac{(\mu_i - 1)\left[\sum_{j=1}^{N}\left(l_{ij}a_{ji}\right)\right]}{\left(1-(\mu_i-1)\left[\sum_{j=1}^{N}\left(l_{ij}a_{ji}\right)\right]\right)} = \frac{1}{\left(1+(1-\mu_i)\left[\sum_{j=1}^{N}\left(l_{ij}a_{ji}\right)\right]\right)} \in [0,1].$$

It follows that the ith row of the matrix inversion is given by

$$\mathbf{e}_i\left(\mathbf{I}-\mathbf{A}\boldsymbol{\mu}\right)^{-1} = \mathbf{e}_i\sum_{n=0}^{\infty}\left(\mathbf{LAG}_{\mu_i}\right)^n\mathbf{L} = \varpi_i\begin{bmatrix} l_{i1} & l_{i2} & \dots & l_{in}\end{bmatrix} = \varpi_i\mathbf{e}_i(\mathbf{I}-\mathbf{A})^{-1} = \varpi_i\mathbf{e}_i\mathbf{L}, \tag{F.2}$$

which is a distortion to the ith row of the original Leontief matrix. Thus

$$\mathbf{e}_i\left(\mathbf{I}-\mathbf{A}\boldsymbol{\mu}\right)^{-1}\left(-\mathbf{I}+\boldsymbol{\alpha\kappa}\right) = \varpi_i\mathbf{e}_i\mathbf{L}\left(-\mathbf{I}+\boldsymbol{\alpha\kappa}\right) = \varpi_i\mathbf{e}_i\left(-\mathbf{L}+\mathbf{1}\boldsymbol{\kappa}\right),$$

where I use the property that $\mathbf{L}\boldsymbol{\alpha}=\mathbf{1}$. Now the volatility of marginal cost is given by

$$\begin{aligned}\mathbb{V}\left(\mathrm{mc}_{it}\right) &= \left\|\mathbf{e}_i\left[\left(\mathbf{I}-\mathbf{A}\boldsymbol{\mu}\right)^{-1}\left(-\mathbf{I}+\boldsymbol{\alpha\kappa}\right)\right]\boldsymbol{\Sigma}_z^{1/2}\right\|^2 \\ &= \left\|\varpi_i\mathbf{e}_i\left(-\mathbf{L}+\mathbf{1}\boldsymbol{\kappa}\right)\boldsymbol{\Sigma}_z^{1/2}\right\|^2 \\ &= \varpi_i^2\sum_{j=1}^{N}\left[\left(\kappa_j - l_{ij}\right)\sigma_j\right]^2 \\ &= \frac{\sum_{j=1}^{N}\left[\left(\kappa_j-l_{ij}\right)\sigma_j\right]^2}{\left(1+(1-\mu_i)\left[\sum_{j=1}^{N}(l_{ij}a_{ji})\right]\right)^2}.\end{aligned}$$

Then the fixed point system (3.6) follows:

$$\mu_i = 1 - \frac{\chi_i}{\theta_i\lambda_i}\frac{\left(1+(1-\mu_i)\left[\sum_{j=1}^{N}(l_{ij}a_{ji})\right]\right)^2}{\sum_{j=1}^{N}\left[\left(\kappa_j-l_{ij}\right)\sigma_j\right]^2} \tag{F.3}$$

([F.3](#)) is a quadratic equation in terms of μ_i,

$$\mathcal{S}_i^2 \bar{x}_i^2 + (2\mathcal{S}_i - \mathcal{N}_i)\, \bar{x}_i + 1 = 0, \tag{F.4}$$

where we define[46]

$$\bar{x}_i = (1 - \mu_i) \in [0, 1]; \qquad \mathcal{S}_i = \left[\sum_{j=1}^{N} (l_{ij} a_{ji})\right] \geq 0; \qquad \mathcal{N}_i = \frac{\theta_i \lambda_i}{\chi_i} \sum_{j=1}^{N} \left[\left(\kappa_j - l_{ij}\right) \sigma_j\right]^2 .$$

We also define $\mathcal{Q}_i = \sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}$ as the discriminant of the above quadratic equation. If $0 < \mathcal{S}_i < 1$ and $\mathcal{N}_i > (1 + \mathcal{S}_i)^2$, it is easy to deduce that

$$-\frac{2\mathcal{S}_i - \mathcal{N}_i}{2\mathcal{S}_i^2} > \frac{(1+\mathcal{S}_i)^2 - 2\mathcal{S}_i}{2\mathcal{S}_i^2} > 1, \qquad \mathcal{Q}_i = \sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i} > 0, \tag{F.5}$$

Therefore, the quadratic curver centers at $\bar{x}_i = -\frac{2\mathcal{S}_i - \mathcal{N}_i}{2\mathcal{S}_i^2} > 1$ with two distinct solutions,

$$\bar{x}_i = \frac{(\mathcal{N}_i - 2\mathcal{S}_i) \pm \sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}}{2\mathcal{S}_i^2} \tag{F.6}$$

It is clear that the bigger solution is larger than 1, so all we need to show is that the smaller solution is bounded above by 1. By the property of quadratic function, it suffices to show that at $\bar{x}_i = 1$, the function $g(\bar{x}_i) \equiv \mathcal{S}_i^2 \bar{x}_i^2 + (2\mathcal{S}_i - \mathcal{N}_i)\, \bar{x}_i + 1$ is negative,

$$g(1) = \mathcal{S}_i^2 + (2\mathcal{S}_i - \mathcal{N}_i) + 1 < 0$$

which is satisfied by our assumption.

Now if $\mathcal{S}_i = 0$ and $\mathcal{N}_i > (1 + \mathcal{S}_i)^2 = 1$, the quadratic equation ([F.4](#)) degenerate into a linear equation in terms of $\bar{x}_i$,

$$\bar{x}_i = \frac{1}{\mathcal{N}_i} \in (0, 1) \tag{F.7}$$

To summarize, we collate the definitions of notations $\bar{x}_i$, $\mathcal{S}_i$ and $\mathcal{N}_i$ in ([F.4](#)), the above conditions show that if $\sum_{j=1}^{N} l_{ij} a_{ji} < 1$ and $\frac{\theta_i \lambda_i}{\chi_i} \sum_{j=1}^{N} \left[\left(\kappa_j - l_{ij}\right) \sigma_j\right]^2 > \left[1 + \sum_{j=1}^{N} \left(l_{ij} a_{ji}\right)\right]^2$, the unique equilibrium solution is given by

$$\mu_i = 1 - \frac{\frac{\theta_i \lambda_i}{\chi_i} \sum_{j=1}^{N} \left[\left(\kappa_j - l_{ij}\right) \sigma_j\right]^2 - 2\sum_{j=1}^{N} \left(l_{ij} a_{ji}\right) - \sqrt{\frac{\theta_i \lambda_i}{\chi_i} \sum_{j=1}^{N} \left[\left(\kappa_j - l_{ij}\right) \sigma_j\right]^2 \left[\frac{\theta_i \lambda_i}{\chi_i} \sum_{j=1}^{N} \left[\left(\kappa_j - l_{ij}\right) \sigma_j\right]^2 - 4\sum_{j=1}^{N} (l_{ij} a_{ji})\right]}}{2\left(\sum_{j=1}^{N} \left(l_{ij} a_{ji}\right)\right)^2}. \tag{F.8}$$

[46] By definition of the Leontief inverse, $\mathcal{S}_i = \left[\sum_{j=1}^{N} (l_{ij} a_{ji})\right] \geq 0$.

We notice that (F.7) and (F.8) coincides at $\mathcal{S}_i = 0$ by L'Hôpital's rule,

$$\lim_{\mathcal{S}_i \to 0} \bar{x}_i = \lim_{\mathcal{S}_i \to 0} \frac{-2 + 2\mathcal{N}_i \left(\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i\right)^{-\frac{1}{2}}}{4\mathcal{S}_i} = \lim_{\mathcal{S}_i \to 0} \frac{4\mathcal{N}_i^2 \left(\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i\right)^{-\frac{3}{2}}}{4} = \frac{1}{\mathcal{N}_i} = \frac{\chi_i}{\theta_i \lambda_i} \frac{1}{\sum_{j=1}^N \left[\left(\kappa_j - l_{ij}\right)\sigma_j\right]^2}. \quad \text{(F.9)}$$

The proof is now complete. The next two Propositions inherit the same parameter restrictions as in Proposition F.2. □

To digest the message of Proposition F.2, we note that the volatity of marginal cost admits representation,

$$\mathbb{V}\left(mc_{it}\right) = \frac{\sum_{j=1}^N \left[\left(\kappa_j - l_{ij}\right)\sigma_j\right]^2}{\left(1 + (1 - \mu_i)\left[\sum_{j=1}^N \left(l_{ij}a_{ji}\right)\right]\right)^2} = \varpi_i^2 \sum_{j=1}^N \left[\left(\kappa_j - l_{ij}\right)\sigma_j\right]^2; \qquad \varpi_i \in [0,1].$$

The term $\sum_{j=1}^N \left[\left(\kappa_j - l_{ij}\right)\sigma_j\right]^2$ captures the upstream uncertainty of sector i's suppliers determined by policy and production network's Leontief inverse. The term $\varpi = \frac{1}{\left(1+(1-\mu_i)\left[\sum_{j=1}^N\left(l_{ij}a_{ji}\right)\right]\right)} \in [0,1]$ captures a dampening factor due to endogenous feedbacks of attentions (within-sector complementarity). The intuition is straightforward: inattentive pricing decisions reduces input (marginal cost) uncertainties via production network, feedbacking to further inattentiveness. This dampening factor vanishes when information is perfect ($\mu_i = 1$), but increases in the nominal rigidities ($\mu_i \downarrow$). It also hinges on the downstream (suppliers → customers) propogation patterns of the production network, operated through i's customer then its suppliers ($a_{ji}l_{ij}$).[47]

Suppose the parameter restrictions in Proposition F.2 are satisfied so that a unique equilibrium exists and is given by (F.1). Then we make two important theoretical predictions.

Proposition F.3. *Sector i's equilibrium attention is increasing in its productivity shock's volatility, $\frac{\partial \mu_i}{\partial \sigma_i^2} > 0$.*

Proof. Recall from equations (F.6) and (F.7), if $\mathcal{S}_i = 0$,

$$\frac{\partial \bar{x}_i}{\partial \mathcal{N}_i} = -\frac{1}{\mathcal{N}_i^2} < 0$$

On the other hand, if $0 < \mathcal{S}_i < 1$,[48]

[47]This feature relies on the assumptions of Cobb-Douglas technology and monopolistic competition. More generally, we note that in matrix form,

$$\sum_{j=1}^N \left(l_{ij}a_{ji}\right) = \left[(I - A)^{-1}A\right]_{(i,i)} = \left[A + A^2 + ...\right]_{(i,i)}.$$

[48]By continuity of derivatives, it is easy to verify using the L'Hôpital's rule that

$$\lim_{\mathcal{S}_i \to 0^+} \frac{\partial \bar{x}_i}{\partial \mathcal{N}_i} = \lim_{\mathcal{S}_i \to 0^+} \frac{1}{2\mathcal{S}_i^2}\left(1 - \frac{(\mathcal{N}_i - 2\mathcal{S}_i)}{\sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}}\right) = -\frac{1}{\mathcal{N}_i^2} < 0.$$

$$
\frac{\partial \bar{x}_i}{\partial \mathcal{N}_i} = \frac{1}{2\mathcal{S}_i^2}\left(1 - \frac{(\mathcal{N}_i - 2\mathcal{S}_i)}{\sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}}\right) < 0,
$$

where $\bar{x}_i = (1-\mu_i)$, $\mathcal{S}_i = \left[\sum_{j=1}^{N}(l_{ij}a_{ji})\right]$ and $\mathcal{N}_i = \frac{\theta_i\lambda_i}{\chi_i}\sum_{j=1}^{N}\left[\left(\kappa_j - l_{ij}\right)\sigma_j\right]^2$. Combining $\frac{\partial \mu_i}{\partial \bar{x}_i} = -1$ and $\frac{\partial \mathcal{N}_i}{\partial \sigma_i^2} = \frac{\theta_i\lambda_i}{\chi_i}(k_i - l_{ii})^2 \geq 0$,

$$
\frac{\partial \mu_i}{\partial \sigma_i^2} = \frac{\partial \mu_i}{\partial \bar{x}_i}\frac{\partial \bar{x}_i}{\partial \mathcal{N}_i}\frac{\partial \mathcal{N}_i}{\partial \sigma_i^2} \geq 0
$$

□

More generally, sectoral attention μ_i increases when firms in this sector face higher upstream uncertainty captured by $\sum_{j=1}^{N}\left[\left(\kappa_j - l_{ij}\right)\sigma_j\right]^2$. The prediction of Proposition F.3 implies a NEGATIVE correlation between sectoral nominal rigidities and shock uncertainties. Such correlation is absent in the Calvo model of sticky prices, the model of exogenous-information frictions, and in models of fixed-capacity information acquisition. In quantitative analysis in Section 5 , we show that the negative correlation continues to hold when all sectors in the network are subjective to rational inattention. In this regard, the above proposition provides a testable empirical prediction that helps distinguish different modeling approaches of information-driven nominal rigidities.

Proposition F.3 characertizes the attention-volatility correlation from the perspective of browser. The next proposition characterize the endogenous feedback effect of attentions from the perspective of browsees.

Proposition F.4. *A firm in sector i has lower attention if sector i's relative importance as a supplier increases. That is,* $\frac{\partial \mu_i}{\partial \sum_{j=1}^{N}(l_{ij}a_{ji})} < 0$.[49]

Proof. Recall from equation (F.4), if $\mathcal{S}_i > 0$

$$
\begin{aligned}
\frac{\partial \bar{x}_i}{\partial \mathcal{S}_i} &= \frac{2\mathcal{S}_i^2\left[-2 + 2\mathcal{N}_i\left(\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i\right)^{-\frac{1}{2}}\right] - 4\mathcal{S}_i\left[(\mathcal{N}_i - 2\mathcal{S}_i) - \sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}\right]}{4\mathcal{S}_i^4} \\
&= \frac{\mathcal{S}_i + \frac{\mathcal{S}_i\mathcal{N}_i}{\mathcal{Q}_i} - \mathcal{N}_i + \mathcal{Q}_i}{\mathcal{S}_i^3} \\
&= \frac{\mathcal{S}_i\mathcal{Q}_i + \mathcal{S}_i\mathcal{N}_i - \mathcal{N}_i\mathcal{Q}_i + \mathcal{Q}_i^2}{\mathcal{S}_i^3\mathcal{Q}_i}.
\end{aligned}
\tag{F.10}
$$

[49] We assume the input certainty $\sum_{j=1}^{N}\left[\left(\kappa_j - l_{ij}\right)\sigma_j\right]^2$ remains unchanged by adjusting the sectoral shock volatilities $\left\{\sigma_j^2\right\}_{j=1}^{N}$.

Using some algebraic tricks, we simplify the equation as

$$\begin{aligned}\mathcal{S}_i\mathcal{Q}_i + \mathcal{S}_i\mathcal{N}_i - \mathcal{N}_i\mathcal{Q}_i + \mathcal{Q}_i^2 &= \mathcal{S}_i\sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i} + \mathcal{S}_i\mathcal{N}_i - \mathcal{N}_i\sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i} + \mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i \\ &= \frac{1}{2}\left(\mathcal{N}_i - \sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}\right)^2 - \mathcal{S}_i\left(\mathcal{N}_i - \sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}\right),\end{aligned}$$

Given the parameter restrictions specified in Proposition F.2, $\frac{\partial \bar{x}_i}{\partial \mathcal{S}_i} > 0$ if and only if $\left(\mathcal{N}_i - \sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}\right) > 2\mathcal{S}_i$. Now construct a continuous function $F(\mathcal{S}_i) = \left(\mathcal{N}_i - \sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}\right) - 2\mathcal{S}_i$ on $\mathcal{S}_i \in [0, 1)$. We notice that

$$\frac{dF(\mathcal{S}_i)}{d\mathcal{S}_i} = \frac{2\mathcal{N}_i}{\sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}} - 2 > 0 \tag{F.11}$$

when $\mathcal{S}_i > 0$. Meanwhile, $F(\mathcal{S}_i)\big|_{\mathcal{S}_i=0} = 0$. Therefore, $F(\mathcal{S}_i)\big|_{\mathcal{S}_i>0} > 0$, implying that $\frac{\partial \bar{x}_i}{\partial \mathcal{S}_i} > 0$ based on conditions (F.10) - (F.11).

When $\mathcal{S}_i = 0$, we notice that by (F.9) and (F.7), the function $\bar{x}_i(\mathcal{S}_i)$ is continuous on $\mathcal{S}_i = 0^+$. Applying the L'Hôpital's rule,

$$\begin{aligned}\frac{\partial \bar{x}_i}{\partial \mathcal{S}_i}\bigg|_{\mathcal{S}_i=0^+} = \lim_{\mathcal{S}_i\to 0^+}\frac{\bar{x}_i(\mathcal{S}_i > 0) - \bar{x}_i(\mathcal{S}_i = 0)}{\mathcal{S}_i - 0} &= \lim_{\mathcal{S}_i\to 0^+}\frac{\frac{(\mathcal{N}_i - 2\mathcal{S}_i) - \sqrt{\mathcal{N}_i^2 - 4\mathcal{N}_i\mathcal{S}_i}}{2\mathcal{S}_i^2} - \frac{1}{\mathcal{N}_i}}{\mathcal{S}_i} \\ &= \lim_{\mathcal{S}_i\to 0^+}\frac{\mathcal{N}_i - 2\mathcal{S}_i - \frac{2\mathcal{S}_i^2}{\mathcal{N}_i} - \mathcal{Q}_i}{2\mathcal{S}_i^3} \\ &= \lim_{\mathcal{S}_i\to 0^+}\frac{-2 - \frac{4\mathcal{S}_i}{\mathcal{N}_i} - \frac{\partial \mathcal{Q}_i}{\partial \mathcal{S}_i}}{6\mathcal{S}_i^2} \\ &= \lim_{\mathcal{S}_i\to 0^+}\frac{-\frac{4}{\mathcal{N}_i} - \frac{\partial^2 \mathcal{Q}_i}{\partial \mathcal{S}_i^2}}{12\mathcal{S}_i} \\ &= \lim_{\mathcal{S}_i\to 0^+}\frac{-\frac{\partial^3 \mathcal{Q}_i}{\partial \mathcal{S}_i^3}}{12} = \frac{2}{\mathcal{N}_i^2} > 0\end{aligned} \tag{F.12}$$

where $\lim_{\mathcal{S}_i\to 0^+}\mathcal{Q}_i = \mathcal{N}_i$, $\lim_{\mathcal{S}_i\to 0^+}\frac{\partial \mathcal{Q}_i}{\partial \mathcal{S}_i} = -2$, $\lim_{\mathcal{S}_i\to 0^+}\frac{\partial^2 \mathcal{Q}_i}{\partial \mathcal{S}_i^2} = -\frac{4}{\mathcal{N}_i}$, and $\lim_{\mathcal{S}_i\to 0^+}\frac{\partial^3 \mathcal{Q}_i}{\partial \mathcal{S}_i^3} = -\frac{24}{\mathcal{N}_i^2}$.

Finally, recall from the definition of $\bar{x}_i = 1 - \mu_i$ and $\mathcal{S}_i = \sum_{j=1}^N (l_{ij}a_{ji})$, we obtain

$$\frac{\partial \mu_i}{\partial \sum_{j=1}^N (l_{ij}a_{ji})} = -\frac{\partial \bar{x}_i}{\partial \mathcal{S}_i} < 0$$

as desired. $\square$

Intuitively, when sector i becomes a more important supplier, each firm (i, ι)'s inattention has a larger impact on other firms because it is being more closely watched as a browsee by other firms in the sector. Therefore, sector i's nominal rigidity has stronger dampening feedback effect on the equilibrium attention of firms in this sector. While our analysis in this example focus on within-sector strategic complementarity, the underlying intuition also applies to more general setting of inattentive production network with multivariate attention linkages and inter-sector strategic complementarities, and we verify this insight in Section 5 .

F.3 Strategic Complementarity of Sectoral Attentions and Nominal Rigidities

Using Proposition 4.2, we define $\mathcal{T}_{\boldsymbol{\mu}^v}$ as the strategic-complementarity matrix of attentions among sectors, capturing attention linkages in the production network. The next corollary shows that sectoral attentions are **indeed** strategic complements when monetary policy completely stabilize wages.

Corollary F.1. *If $\boldsymbol{\kappa} = 0$, the strategic complementarity matrix of attentions $\mathcal{T}_{\boldsymbol{\mu}^v} \geq 0$ is a non-negative matrix.*

Proof. Recall an expression of $\mathcal{T}_{\boldsymbol{\mu}^v}$ in equation (E.30), if $\boldsymbol{\kappa} = \mathbf{0}$, then $\boldsymbol{\Gamma} = (-\mathbf{I} + \boldsymbol{\alpha}\kappa) = -\mathbf{I}$ and

$$\mathcal{T}_{\boldsymbol{\mu}^v} = 2\,\mathrm{diag}\left(\left\{\frac{\chi_i}{\theta_i \lambda_i V_i^2}\right\}_{i=1}^N\right)\left[\left(\boldsymbol{\Delta}_\mu \boldsymbol{\Gamma}\boldsymbol{\Sigma}_z \boldsymbol{\Gamma}'\boldsymbol{\Delta}'_\mu\right) \odot \left(\boldsymbol{\Delta}_\mu \mathbf{A}\right)\right] = 2\,\mathrm{diag}\left(\left\{\frac{\chi_i}{\theta_i \lambda_i V_i^2}\right\}_{i=1}^N\right)\left[\left(\boldsymbol{\Delta}_\mu \boldsymbol{\Sigma}_z \boldsymbol{\Delta}'_\mu\right) \odot \left(\boldsymbol{\Delta}_\mu \mathbf{A}\right)\right] \geq 0, \quad \text{(F.13)}$$

since $\chi_i, \theta_i, \lambda_i > 0$ by definition, and $\boldsymbol{\Delta}_\mu > 0$ from (D.41). $\square$

If we express the feedback (complementarity) matrix as $\mathcal{T}_{\boldsymbol{\mu}^v} = \begin{bmatrix} \frac{\partial \mathcal{T}_1}{\partial \mu_1} & \frac{\partial \mathcal{T}_1}{\partial \mu_2} & \cdots & \frac{\partial \mathcal{T}_1}{\partial \mu_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathcal{T}_N}{\partial \mu_1} & \frac{\partial \mathcal{T}_N}{\partial \mu_2} & \cdots & \frac{\partial \mathcal{T}_N}{\partial \mu_N} \end{bmatrix}$, the diagonal entries capture the WITHIN-sector feedbacks while off-diagonal elements represent the INTER-sector feedbacks of attentions. Clearly, the these strategic linkages are jointly controlled by the triple $(\mathbf{A}, \boldsymbol{\Sigma}_{\boldsymbol{z}}, \boldsymbol{\kappa})$.

G. Calibration and Quantitative Analysis

G.1 Calibration of Information Cost

We use indirect inference to calibrate the marginal information cost χ_i in different sectors. In our models with monetary policy, any model-implied moment is endogeous/variant to the choice of policy. Therefore, we first fix a policy environment that is consistent with the conduct of monetary policy in our data sample period.

CPI policy environment. We focus on a model with monetary policy that targets CPI stability (inflation targeting). This is consistent with the emprical evidence that within our sample period, the CPI is largely constant. We consider the following price stabilization

$$\text{CPI} = \sum_{i=1}^{N} \varphi_i p_{it} = 0; \qquad \varphi_i = \beta_i; \qquad \forall i = 1, 2, \ldots N \tag{G.1}$$

where $\boldsymbol{\beta}$ is calibrated to match the share of final uses of each industry's output. La'O and Tahbaz-Salehi (2022) argue that this formulation is akin to CPI stabilization policy.

Monetary Policy Rule. Price stabilization (G.1) is an implication and feature of the monetary policy, but not the policy itself. So next, we convert the price stabilization into the corresponding monetary policy rule. Using the equilibrium pricing function (3.7), we find that

$$\sum_{i=1}^{N} \varphi_i p_{it} = \boldsymbol{\varphi}(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}(-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa})\boldsymbol{z}_t = 0, \tag{G.2}$$

which holds for any productivity shock vector $\boldsymbol{z}_t$. Thus, $\boldsymbol{\varphi}(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}(-\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\kappa}) \equiv 0$ such that

$$\boldsymbol{\kappa} = \frac{1}{\boldsymbol{\varphi}'(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}\boldsymbol{\alpha}} \boldsymbol{\varphi}'(\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}, \tag{G.3}$$

which provides a correspondence between the monetary policy (wage) rule $\boldsymbol{\kappa}$ and the price stabilization policy weight $\boldsymbol{\varphi}$. As Lemma E.3 predicts, any price-stabilization policy satisfies the condition $\boldsymbol{\kappa}\boldsymbol{\alpha} = 1$.

Equilibrium under CPI stability. Under CPI policy environment, the general equilibrium fixed-point is determined by the equation system (3.6), which implicitly define the mapping from CPI-based $\boldsymbol{\kappa}^{CPI}$ to $\boldsymbol{\mu}^{CPI}$ as

$$\boldsymbol{\mu}^{CPI} = \mathcal{F}\left(\boldsymbol{\kappa}^{CPI}\right) \tag{G.4}$$

Meanwhile, to achieve CPI stability, the monetary policy rule is also endogenous, given by (G.3) and $\boldsymbol{\varphi} = \boldsymbol{\beta}$, we have

$$\left(\boldsymbol{\kappa}^{CPI}\right)' = \frac{1}{\boldsymbol{\beta}'(\mathbf{I} - \boldsymbol{\mu}^{CPI}\mathbf{A})^{-1}\boldsymbol{\mu}^{CPI}\boldsymbol{\alpha}} \boldsymbol{\beta}'(\mathbf{I} - \boldsymbol{\mu}^{CPI}\mathbf{A})^{-1}\boldsymbol{\mu}^{CPI}. \tag{G.5}$$

Therefore, the fixed point system (G.4) and (G.5) jointly determines equilibrium under CPI stability.

Forecast error in model. We first derive the expression for firms' revenue. Recall that The output of firm ι in sector i is determined by $Y_{i\iota t} = (P_{i\iota t}/P_{it})^{-\theta_i} Y_{it}$. The demand of firms in sector j for intermediate goods produced in sector i is given by

$$\begin{aligned} X_{jit} &= \int_0^1 X_{j\iota it} d\iota \\ &= a_{ji} \frac{\mathrm{MC}_{jt}}{P_{it}} \int_0^1 Y_{j\iota t} d\iota \\ &= a_{ji} \frac{P_{jt} Y_{jt}}{P_{it}} \mathrm{MC}_{jt} P_{jt}^{\theta_j - 1} \int_0^1 P_{j\iota t}^{-\theta_j} d\iota \\ &= a_{ji} \frac{P_{jt} Y_{jt}}{P_{it}} \gamma_{jt}, \end{aligned}$$

where $\gamma_{jt} = MC_{jt} P_{jt}^{\theta_j - 1} \int_0^1 P_{j\iota t}^{-\theta_j} d\iota$ is a function of prices.

The market clearing condition for sector good i is given by

$$Y_{it} = C_{it} + \sum_{j=1}^{N} \int_0^1 X_{j\iota it} d\iota.$$

which implies that

$$P_{it} Y_{it} = \beta_i P_t C_t + \sum_{j=1}^{N} a_{ji} P_{jt} Y_{jt} \gamma_{jt}$$

Because the aggregate nominal spending is equal to the money supply $P_t C_t = M_t$, we have

$$P_{it} Y_{it} = \beta_i M_t + \sum_{j=1}^{N} a_{ji} P_{jt} Y_{jt} \gamma_{jt}$$

Log-linearizing the above equation around the steady state yields

$$\lambda_i (p_{it} + y_{it}) = \beta_i m_t + \sum_{j=1}^{N} a_{ji} \lambda_j (y_{jt} + \sum_{k=1}^{N} a_{jk} p_{kt} - z_{jt})$$

We derive the expression for sectoral outputs in vector form,

$$\boldsymbol{y}_t = \left[(\mathbf{I} - \mathbf{A}') \operatorname{diag}(\boldsymbol{\lambda}) \right]^{-1} \left[\boldsymbol{\beta} m_t + \left(\mathbf{A}' \operatorname{diag}(\boldsymbol{\lambda}) \mathbf{A} - \operatorname{diag}(\boldsymbol{\lambda}) \right) \boldsymbol{p}_t - \mathbf{A}' \operatorname{diag}(\boldsymbol{\lambda}) \boldsymbol{z}_t \right] \tag{G.6}$$

where the matrix form of monetary supply m_t is given by Proposition 3.6,

$$m_t = \boldsymbol{\psi}' \boldsymbol{z}_t = \frac{\eta}{1 + \gamma \eta} \boldsymbol{\kappa} \boldsymbol{z}_t + \left(1 - \frac{\eta}{1 + \gamma \eta} \right) \boldsymbol{\beta}' (\mathbf{I} - \boldsymbol{\mu} \mathbf{A})^{-1} \boldsymbol{\mu} (-\mathbf{I} + \boldsymbol{\alpha} \boldsymbol{\kappa}) \boldsymbol{z}_t + \frac{1}{1 + \gamma \eta} \boldsymbol{\lambda}' \boldsymbol{z}_t.$$

ψ is $N \times 1$ column vector and $\boldsymbol{p}_t = \phi \boldsymbol{z}_t = (\mathbf{I} - \boldsymbol{\mu}\mathbf{A})^{-1}\boldsymbol{\mu}(-\mathbf{I} + \boldsymbol{\alpha\kappa})\boldsymbol{z}_t$. Substituting the expression of $\boldsymbol{p}_t$ and m_t into equation (G.6) we obtain the sectoral revenue as a function of fundamental shocks $\boldsymbol{z}_t$,

$$\boldsymbol{r}_t = \boldsymbol{p}_t + \boldsymbol{y}_t = \boldsymbol{\Psi}_r \boldsymbol{z}_t,$$

where

$$\boldsymbol{\Psi}_r = \left[\mathbf{I} + \left[(I - \mathbf{A}')\,\mathrm{diag}(\boldsymbol{\lambda}) \right]^{-1} \left(\mathbf{A}'\,\mathrm{diag}(\boldsymbol{\lambda})\mathbf{A} - \mathrm{diag}(\boldsymbol{\lambda}) \right) \right] \phi + \left[(I - \mathbf{A}')\,\mathrm{diag}(\boldsymbol{\lambda}) \right]^{-1} \left[\beta\boldsymbol{\psi}' - \mathbf{A}'\,\mathrm{diag}(\boldsymbol{\lambda}) \right] \tag{G.7}$$

The revenue forecast error of i-sector is given by

$$\mathrm{FE}^{rev}_{i,t} = \frac{\left| \mathbf{e}_i \left[\boldsymbol{\Psi}_r \left(\mathbf{I} - (\mathbf{I} - \boldsymbol{\Sigma}_{z|x_i} \boldsymbol{\Sigma}_z^{-1}) \right) \boldsymbol{z}_t \right] \right|}{\left| \mathbf{e}_i \boldsymbol{\Psi}_r \boldsymbol{z}_t + 1 \right|} = \frac{\left| \mathbf{e}_i \boldsymbol{\Psi}_r \boldsymbol{\Sigma}_{z|x_i} \boldsymbol{\Sigma}_z^{-1} \boldsymbol{z}_t \right|}{\left| \mathbf{e}_i \boldsymbol{\Psi}_r \boldsymbol{z}_t + 1 \right|}; \qquad \forall i = 1, 2, \ldots N$$

where $\boldsymbol{\Sigma}_{z|x_i}$ is the posterior covariance matrices of the fundamental shocks $\boldsymbol{z}_t$ from Lemma D.1.

The firm ι's profit is given by

$$\Pi_{i\iota t} = (1 + \tau_i) P_{i\iota t} Y_{i\iota t} - W_t L_{i\iota t} - \sum_{j=1}^{N} P_j X_{i\iota j t} - T_{i\iota t} = (1 + \tau_i) P_{i\iota t} Y_{i\iota t} - \mathrm{MC}_{it} Y_{i\iota t} - T_{i\iota t}, \tag{G.8}$$

Next, the sectoral level profit before tax becomes

$$\begin{aligned} \Pi^b_{it} = \int (\Pi_{i\iota t} + T_{i\iota t}) &= \int \left[(1 + \tau_i) P_{i\iota t} - \mathrm{MC}_{it} \right] Y_{i\iota t} \\ &= (1 + \tau_i) P_{it} Y_{it} - P_{it} Y_{it} \frac{\mathrm{MC}_{it}}{P_{it}} \int \left(\frac{P_{i\iota t}}{P_{it}} \right)^{-\theta_i} \\ &= P_{it} Y_{it} (1 + \tau_i - \mathcal{E}_{it}) \end{aligned} \tag{G.9}$$

where

$$\mathcal{E}_{it} = \frac{\mathrm{MC}_{it}}{P_{it}} \int \left(\frac{P_{i\iota t}}{P_{it}} \right)^{-\theta_i} \tag{G.10}$$

and its log-linearized version is

$$\varepsilon_{it} = \mathrm{mc}_{it} - p_{it} - \theta_i \int (p_{i\iota t} - p_{it}) = \mathrm{mc}_{it} - p_{it}$$

Also, we have $\mathrm{mc}_{it} = \mu_{it}^{-1} p_{it}$. Thus the log-linearized version of Π_{it}^{b} is

$$\begin{aligned}
\pi_{it}^{b} &= p_{it} + y_{it} + \frac{-\mathcal{E}_i \varepsilon_{it}}{1 + \tau_i - \mathcal{E}_i} \\
&= p_{it} + y_{it} - \frac{\varepsilon_{it}}{\tau_i} \\
&= p_{it} + y_{it} + \frac{p_{it}\left(1 - \mu_{it}^{-1}\right)}{\tau_i}
\end{aligned} \tag{G.11}$$

In matrix form,

$$\boldsymbol{\pi}_t^b = \boldsymbol{p}_t + \boldsymbol{y}_t + \mathrm{diag}\left(\boldsymbol{\tau}\right)^{-1}\left(I - \boldsymbol{\mu}^{-1}\right)\boldsymbol{p}_t = \left[\boldsymbol{\Psi}_r + \mathrm{diag}\left(\boldsymbol{\tau}\right)^{-1}\left(I - \boldsymbol{\mu}^{-1}\right)\boldsymbol{\Psi}_p\right]\boldsymbol{z}_t = \boldsymbol{\Psi}_{eps}\boldsymbol{z}_t \tag{G.12}$$

Similarly, the EPS forecast error of each sector is given by

$$\mathrm{FE}_{i,t}^{eps} = \frac{\left|\mathbf{e}_i \boldsymbol{\Psi}_{eps} \boldsymbol{\Sigma}_{z|x_i} \boldsymbol{\Sigma}_z^{-1} \boldsymbol{z}_t\right|}{\left|\mathbf{e}_i \boldsymbol{\Psi}_{eps} \boldsymbol{z}_t + 1\right|}; \qquad \forall i = 1, 2, \ldots N$$

Calibration strategy. Given the equilibrium system system (G.4) and (G.5) under CPI stabilization policy, we simulate different EPS forecast error when we calibrate the information cost χ. In particular, we impose the following functional form

$$\log \chi_i = \delta_0 + \delta_1 \log \lambda_i + \delta_2 \log \sigma_i.$$

We calibrate the parameters δ_0, δ_1 and δ_2 to match the distribution of *adjusted* forecast errors of earnings per share (EPS) at the sector level, which are directly related to the informational frictions. Specifically, we choose δ_0, δ_1 and δ_2 to minimize the following loss function,

$$\min_{\delta_0, \delta_1, \delta_2} \left(\mathrm{FE}_{i=25\%}^{eps,data} - \mathrm{FE}_{i=25\%}^{eps,model}\right)^2 + \left(\sum_{i=1}^{N} \frac{1}{N} \mathrm{FE}_i^{eps,data} - \sum_{i=1}^{N} \frac{1}{N} \mathrm{FE}_i^{eps,model}\right)^2 + \left(\mathrm{FE}_{i=75\%}^{eps,data} - \mathrm{FE}_{i=75\%}^{eps,model}\right)^2,$$

where FE_i is the mean absolute forecast error of sector i . We target the mean level of the forecast error, the 25 percentile and 75 percentile of the forecast error in the sectoral distribution. However, the volatility of EPS differs in the data and in the model. To ensure that the forecast error is comparable, we normalize it by the volatility of EPS, both in the data and in the model.

G.2 Computation

Numerical Algorithms. We provide two algorithms for solving the optimal policy problem defined in Definition 4. The first method is based on the first-oder condition (FOC). The second method utilizes constrained nonlinear programming.

FOC Method. We can solve the optimal monetary policy problem by the FOC condition:

$$\frac{dL}{d\boldsymbol{\kappa}} = 0,$$

which finally yields the equilibrium system in Proposition E.2; that is, the jointly fixed-point system (E.7) and (E.65) with respect to variables $\boldsymbol{\mu}^v$ and $\boldsymbol{\kappa}$. We employ Matlab's toolbox 'fsolve' to solve this static and nonlinear system. When the unique solution to this system exists, the optimal monetary policy achieves price stabilization simultaneously by Proposition 4.3. It is worth noting that $\mu_i > 0$ for all sectors $i = 1, 2, ...N$ is implicit in the RI condition (D.26), and the solution of $\boldsymbol{\mu}^v$ to FOC must be an interior solution. The FOC method does not apply when the solution of the optimization problem is on the zero boundary of $\boldsymbol{\mu}^v$. In this case of corner solution, we can awalys use the following constrained nonlinear programming algorithm to solve the problem.

Nonlinear Constrained Programming Method. The optimal monetary policy problem defined in Definition 4 can be transformed into a nonlinear programming problem and restated as

$$\max_{\varrho} \quad L = \Delta U_t^{within} + \Delta U_t^{OG} + \Delta U_t^{across}$$

subject to

$$\varrho = \begin{bmatrix} \boldsymbol{\mu}^v \\ \boldsymbol{\kappa} \end{bmatrix}$$

and nonlinear constraint (E.7). The nonlinear constrained programming (NP) method nests the FOC method as a special case. It handles both the interior solution and the corner solution. That is, the NP algorithm allows for the corner solution case where $\exists \mu_i = 0$, $i = 1, 2, ...N$. When an optimal solution under the FOC structure exists, the NP algorithm will also deliver this interior solution that coincides with the FOC solution. Numerically, we use Matlab's toolbox 'fmincon' to solve the NP problem.

Accuracy of Solution. Utilizing the NP algorithm, we test the accuracy and correctness of our FOC-based characterization of the optimal monetary policy in Section E. Specifically, under the calibration in Table 6, the FOC-based optimal policy solution and the solution based on NP algorithm coincide. Therefore, under our model calibration, an interior optimal policy exists and is given by our characterization in Section E. Figure G.1 plots the optimal solution in terms of $\boldsymbol{\mu}$ and $\boldsymbol{\varphi}$ using two algorithms. The numerical difference of the two solutions in terms of welfare loss is smaller than $1e^{-10}$.

Finally, we find that under model calibration, the OG optimal policy satisfies the condition imposed in Proposition 4.4, that is , $\mathbb{V}\left(c_t - c_t^*\right) = 0$. Figure G.2 plots the output-gap volatility under perturbation around different sectors' optimal OG policy, which shows that the output-gap volatility reduces to 0 at the optimum.

Figure G.1: Accuracy of Solution for Optimal Monetary Policy

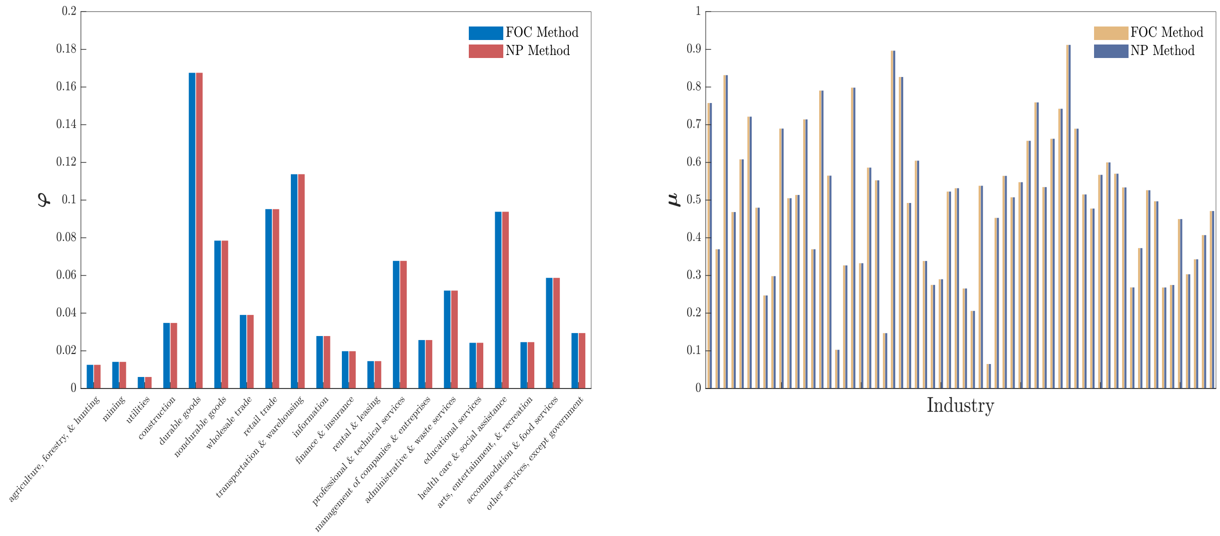


Figure G.2: Completely Stabilized OG Volatility