

Nominal Rigidities, Rational Inattention, and the Optimal Monetary Policy*

Shengliang Ou[†] Penghui Yin[‡] Donghai Zhang[§] Renbin Zhang[¶]

July 10, 2025

Abstract

Optimal monetary policy has traditionally assigned *greater* importance to stabilizing prices in sectors with stickier prices, based on multi-sector models assuming full information or exogenous information frictions. This paper challenges the prevailing policy prescription by introducing rational inattention with endogenous information acquisition. Surprisingly, the optimal policy assigns a *smaller* weight to sectors with stickier prices when the cost of information acquisition is sufficiently high. This counterintuitive result arises from the *endogenous* relationship between firms' attention and nominal rigidities: firms in sectors with more flexible prices pay less attention to macroeconomic conditions. We provide empirical evidence supporting this mechanism.

Keywords: Sticky Prices, Endogenous Information Acquisition, Optimal Monetary Policy, Price Index Stabilization

JEL Classification: E31, E32, E52, E58

*A previous version of this paper was circulated under the title: "Nominal Rigidities, Rational Inattention, and the Optimal Price Index Stabilization Policy". We thank Kosuke Aoki, Pierpaolo Benigno, Alexandre Kohlhas, Minghao Li, Taisuke Nakata, Mirko Wiederholt, and Jieran Wu, as well as the seminar participants at Tokyo University, Keio University, Peking University, Queen Mary University, Xiamen University, UIBE, and Shandong University for insightful discussions. Shengliang Ou thanks the National Natural Science Foundation of China for financial support through Grant 72103121, and the Chenguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission through Grant 22CGA40. Penghui Yin thanks the Natural Science Foundation of Shandong Province under grant ZR2022QG043 and the National Natural Science Foundation of China for financial support through Grant 72303129. Donghai Zhang thanks Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for financial support under grant 495072044. Renbin Zhang thanks the National Natural Science Foundation of China for financial support through Grant 72203130. We thank José Lores and Ruzheng Tian for outstanding research assistance and Joseph Hoon for proofreading.

[†]Shanghai Jiao Tong University, shengliang.ou@gmail.com

[‡]Central University of Finance and Economics, yinpenghui2008@gmail.com

[§]Corresponding author. National University of Singapore, donghai.d.zhang@gmail.com. Address: 1 Arts Link, AS2 06-06 Singapore 117570.

[¶]Shandong University, zhang.renbin.ken@gmail.com

1 Introduction

High inflation, such as the recent episode associated with the COVID-19 crisis, harms social welfare. There is a consensus among policymakers and academic researchers that the stabilization of inflation should be one of the primary mandates of central banks. However, there remains a debate about which inflation index central banks should stabilize. Within the policy circle, policymakers have been tracking the Consumer Price Index (CPI), which is defined as the inflation index weighted by consumption weights (or sizes of sectors). Academic researchers, however, have argued that the optimal price index should account for sectoral heterogeneities, such as the different frequencies of price adjustment observed in the data (Aoki 2001, Benigno 2004, and Mankiw and Reis 2003). Specifically, these papers advocate a price index that attaches a *larger* weight to a sector with stickier prices. We refer to this index as the Stickiness-Adjusted Price Index (SPI).

This paper studies the optimal price index (OPI) stabilization policy. In contrast to conventional wisdom, our model suggests an OPI that assigns a *smaller* weight to a sector with stickier prices when information acquisition is sufficiently costly. The SPI stabilization policy results in a greater welfare loss compared to simple CPI stabilization. The key to our findings is the *endogenous* relationship between information acquisition and nominal rigidities: firms in sectors with more flexible prices pay less attention to macroeconomic conditions. We provide direct empirical evidence that supports this mechanism.

In more detail, we analyze the OPI stabilization policy in a multi-sector economy with nominal rigidities and rationally inattentive firms. Monopolistically competitive firms that operate and set prices in different sectors are subject to unobserved sector-specific supply shocks. Firms observe noisy signals about the state (sectoral supply shocks) of the economy. These signals are idiosyncratic, which gives rise to dispersed beliefs. Importantly, following the rational inattention literature (Sims 2003 and Mackowiak and Wiederholt 2009), the precision of the signal is an endogenous choice: firms choose the optimal amount of information by balancing the cost of processing information with its benefit. Firms are subject to nominal rigidities *à la* Calvo in addition to information frictions. Specifically, firms are allowed to reset their prices optimally with a probability less than one. As a result, ex-post, there is a fraction of firms that are staggered with their initial prices. The degree of nominal rigidities — the probability of resetting prices (or equivalently the fraction of price-staggered firms) — is heterogeneous across sectors.

These two frictions give rise to price dispersion, which in turn reduces social welfare. First, dispersed beliefs lead to price dispersions among price-resetting firms. This

is called the dispersed belief price dispersion component. Second, there is price dispersion across the types of firms: price-staggered and price-resetting firms. This is called the Calvo price dispersion component. This paper highlights the interaction between these two frictions, each of which has been shown to be empirically relevant. For instance, Nakamura and Steinsson (2008) and Gorodnichenko and Weber (2016) document heterogeneous *infrequent* price adjustments across sectors, while Coibion and Gorodnichenko (2015) and Andrade et al. (2016) examine information frictions and the dispersed beliefs that can emerge due to rational inattention.¹

We emphasize the importance of jointly considering nominal rigidities and endogenous information choice. Proposition 1 shows that firms in a sector with more flexible prices pay less attention to macroeconomic conditions. This result suggests that sectors with fewer nominal frictions are associated with more severe information frictions due to firms' endogenous choices. The intuition for the interaction of the two frictions is as follows. The uncertainty about whether a firm can reset its price is resolved first. That is, before acquiring information, firms know whether they can adjust prices. In response to a shock, price-resetting firms in a sector with more flexible prices adjust prices to a lesser extent. Intuitively, when productivity increases, price-resetting firms face two opposing forces. On one hand, higher productivity directly lowers marginal costs and pushes prices down. On the other hand, it raises aggregate demand and nominal wages, which partially offsets this price decline. The degree of nominal rigidity determines how strongly the second, general-equilibrium channel counteracts the first: with more rigid prices, wages rise less, and firms end up adjusting prices more in response to shocks.

Firms' responsiveness to shocks is proportional to the marginal benefit of acquiring information. In the extreme case when firms' desired prices react very little to the unobserved state of the economy, the benefit of acquiring information is limited. Therefore, greater price flexibility in a sector reduces the incentive for price-resetting firms in that sector to acquire information.

Before presenting the main results of the paper, it is important to clarify the terminology related to optimal policy. In our baseline model, the OPI stabilization policy is equivalent to the optimal Ramsey policy (Lemma 3.1). Therefore, we will use the terms

¹Note that nominal rigidities in this paper refer to infrequent price adjustments. This is different from a separate literature where nominal frictions arise due to information frictions; see, for example, Angeletos et al. (2016) and Mankiw and Reis (2002). The joint consideration of information frictions and infrequent price adjustment frictions exists in the literature; see, e.g., Nimark (2008) and Ou et al. (2021) for different purposes.

“OPI stabilization” and “optimal monetary policy” interchangeably throughout the baseline analysis.

The model prescribes an optimal monetary policy that challenges conventional wisdom due to the endogenous relationship between nominal frictions and information choice. The baseline model can be adapted to feature full information and exogenous information frictions with minimal modifications. We illustrate that conventional wisdom holds in a model with full information (Benigno 2004) or exogenous information frictions (Ou et al. 2021): the OPI attaches a *larger* weight to a sector with stickier prices. In contrast, when the amount of information firms acquire is an endogenous choice, Proposition 3 demonstrates that the OPI can be altered *qualitatively*: it attaches a *smaller* weight to a sector with stickier prices when information acquisition is sufficiently costly.

The dispersed belief price dispersion component of social welfare drives this result. A decomposition exercise demonstrates that, according to the other components of the social welfare loss function,² the OPI assigns a *larger* weight to a sector with stickier prices. In contrast, when information is an endogenous choice, the dispersed belief price dispersion component suggests that the OPI assigns a *smaller* weight to a sector with stickier prices. Since the total welfare loss function consists of both the dispersed belief price dispersion and the other components, the OPI depends on which channel dominates. This explains *the condition* attached to Proposition 3. Specifically, since the dispersed belief price dispersion component relies on information frictions, the channel we emphasize can overturn conventional wisdom when the cost of information acquisition exceeds a *threshold*. Under standard calibrations of the model parameters, the condition required to overturn conventional wisdom is satisfied.

An interesting corollary arises when the cost of information acquisition coincides with the aforementioned threshold, a condition we label the *Divine Coincidence Condition*. Under the Divine Coincidence Condition, the dispersed belief price dispersion component coincidentally offsets the other channels. As a result, the CPI stabilization policy becomes optimal, *independent* of heterogeneities in nominal rigidities. Therefore, the question, “Is the stabilization of CPI a good policy?” becomes a quantitative one—it depends on how far the calibrated model is from the Divine Coincidence Condition. Using the model as a laboratory, we evaluate the performance of CPI stabilization. As a comparison, the performance of SPI stabilization is also evaluated. The SPI corresponds to the optimal price

²Namely, the Calvo price dispersion, the output gap, and the relative price gap across sector components.

index computed in a model that assumes full information, but which is otherwise equivalent. Surprisingly, CPI stabilization is associated with a smaller social welfare loss in our model than SPI stabilization. Moreover, the paper also discusses how monetary policy shapes firms' endogenous attention choices.

The above results are derived from a static model, which allows for analytical results that are valuable in establishing the qualitative channel we emphasize. The paper proceeds to demonstrate that the main findings carry over to a dynamic setting. In the dynamic model, firms' *dynamic* pricing and information acquisition decisions take into account future profits and the probability of not being able to adjust prices. The mechanism we emphasize remains clearly present in the dynamic model. Extending our analysis to this dynamic setting shows that the calibrated model yields an OPI that assigns a *smaller* weight to a sector with stickier prices. We interpret this as evidence suggesting that the channel we emphasize is quantitatively relevant.

Finally, we present empirical evidence that supports the key mechanism of the paper. Following [Song and Stern \(2022\)](#) and [Flynn and Sastry \(2022\)](#), we use the frequency of macroeconomic keywords in firms' annual reports as a proxy for firms' attention to macroeconomic conditions. To this end, we analyze approximately 200,000 annual reports of U.S. public companies (Form 10-K) and compute the frequency of macroeconomic keywords by dividing the number of occurrences by the total number of words in each report. The firm-level attention data is aggregated to the NAICS 6-digit level to study its correlation with the frequency of price adjustments constructed by [Pasten, Schoenle, and Weber \(2020\)](#). The data confirm the key mechanism highlighted in Proposition 1: firms in sectors with more flexible prices pay less attention to macroeconomic conditions.

Literature Review Our paper highlights the interaction between nominal frictions and endogenous information choice, which has important policy implications, as it qualitatively alters the optimal price index stabilization policy. In doing so, the paper contributes to four strands of literature.

The first strand of literature examines the optimal price index stabilization policy. [Aoki \(2001\)](#), [Mankiw and Reis \(2003\)](#), and [Benigno \(2004\)](#) explore the implications of heterogeneity in the degree of nominal rigidity for the design of an optimal price index. Subsequent works introduce additional features to revisit the optimal price index, such as the input-output structure ([Huang and Liu 2005](#), [Rubbo 2022](#), and [La'O and Tahbaz-Salehi 2022](#)), incomplete financial markets ([Anand, Prasad, and Zhang 2015](#)), capital accumulation ([Basu and Leo 2016](#)), exogenous information frictions ([Ou et al. 2021](#)), and household

heterogeneity (Olivi et al. 2024 and Lan et al. 2024). In our paper, the newly introduced mechanism leads to *qualitatively* different policy implications.

The second strand of literature investigates the implications of rational inattention for the conduct of monetary policy. For example, Mackowiak and Wiederholt (2009), Afrouzi and Yang (2021), Yang (2022), and Afrouzi et al. (2024) examine the implications of rational inattention for the effects of monetary policy. Paciello and Wiederholt (2014) and Li and Wu (2016) investigate the design of optimal monetary policy. In particular, Paciello and Wiederholt (2014) studies a *one-sector* economy with *flexible prices* and finds that a strict price stabilization policy is optimal in response to markup shocks if price-setting firms are rationally inattentive.³ We consider a *multi-sector* setting with *sticky prices*, emphasizing the endogenous relationship between nominal rigidities and information frictions. Specifically, the multi-sector feature of the model allows us to investigate firms' heterogeneous information choices resulting from heterogeneous nominal rigidities, ultimately leading to surprising policy recommendations. Moreover, we contribute to the literature by providing a characterization of the optimal monetary policy in our setting. In a recent paper related to ours, Fang et al. (2024) examines firms' attention choices within a production network setting that abstracts from nominal rigidities and studies the implications for optimal monetary policy.

The third strand of literature emphasizes the interactions between nominal frictions and information frictions. Nimark (2008) demonstrates how aggregate inflation responds gradually and with inertia to shocks due to information frictions. Melosi (2016) and Zhang (2022) highlight that information asymmetries between the central bank and firms (or households) give rise to the signaling channel of monetary policy, thereby altering its transmission. Coibion and Gorodnichenko (2011) provides empirical evidence showing that nominal and information frictions account for over 80% of firms in their estimated model. Ou et al. (2021) shows that increased price flexibility can harm social welfare in the presence of information frictions. Ou et al. (2022) finds that an opaque central bank

³Our paper focuses on firms' endogenous information choice and firm heterogeneities. It is related to a parallel literature that studies household information choice and heterogeneity: see, for example, Broer, Kohlhas, Mitman, and Schlafmann (2022), Gaballo (2016), Luo et al. (2017), Tutino (2013), Yin (2021), and Luo, Nie, and Yin (2022). More broadly, it connects to literature examining the implications of rational inattention for macroeconomics following the seminal work of Sims (2003): see, e.g., Luo (2008), Van Nieuwerburgh and Veldkamp (2010), Mondria (2010), Matějka and McKay (2012), Maćkowiak and Wiederholt (2015), Pasten and Schoenle (2016), Matějka (2016), Stevens (2020), Zorn (2020), Ellison and Macaulay (2021), Miao, Wu, and Young (2022), Miao and Zhang (2024), Ilut and Valchev (2023), and Turen (2023). See also Maćkowiak, Matějka, and Wiederholt (2023) for a literature review and Veldkamp (2011) for a textbook treatment of the topic.

can be optimal when nominal frictions give rise to two types of firms: price-resetting and price-staggered firms. More recently, [Camous and Matveev \(2023\)](#) examines the central bank’s strategic communication by extending these settings to include the possibility of central bank misreporting. Our paper contributes to this literature by extending the framework—and, consequently, the policy analysis—to allow for an intriguing endogenous relationship between the two frictions. In a contemporaneous and independently developed study, [Afrouzi et al. \(2024\)](#) investigates the same endogenous relationship but focuses on its implications for monetary non-neutrality—a different emphasis.

The fourth strand of literature highlights the implications of sectoral heterogeneities for the effects of monetary policy: see, for example, [Alvarez, Lippi, and Oskolkov \(2022\)](#), [Carvalho \(2006\)](#), [Carvalho and Schwartzman \(2015\)](#), [Carvalho, Dam, and Lee \(2020\)](#), [Gautier and Bihan \(2018\)](#), [Höynck, Li, and Zhang \(2022\)](#), [Nakamura and Steinsson \(2008\)](#), and [Pasten, Schoenle, and Weber \(2020\)](#). Our paper illustrates that exogenous sectoral heterogeneities can lead to endogenous sectoral heterogeneities in the degree of information frictions, which have relevant policy implications.

The remainder of this paper is organized as follows: Section 2 introduces the baseline model. Section 3 presents the main results. Section 4 extends the core findings to a dynamic model. Section 5 provides empirical evidence supporting the key mechanism. Finally, Section 6 concludes.

2 The Baseline Model

The model consists of three types of agents: a representative household, a central bank, and firms. The household decides on consumption and labor supply, with the optimization of consumption giving rise to the demand curve faced by firms. Labor is the sole factor of production for firms. The central bank conducts monetary policy.

Throughout the paper, we consider a two-sector economy. To highlight how endogenous information acquisition alters conventional policy prescriptions, we deliberately focus on one specific ex-ante heterogeneity across the two sectors: nominal rigidities. Other sources of heterogeneity, such as sector sizes and shocks’ processes, are excluded. Specifically, we assume that the two sectors are of equal size, each representing half of the economy, and are subject to shocks with the same volatilities.

2.1 Household

The representative household derives utility from consumption (C) and labor supply (L). The utility function is given by: $U(C, \{L_k\}) = \log(C) - \sum_{k=1}^2 L_k$, where L_k is the labor supply to sector k . The consumer's income comes from various sources, including wages earned from labor supply supplied to each sector ($W_k L_k$), profits from firms ($\sum_{k=1}^2 \Pi_k$), and lump-sum transfers or taxes received from the government (T). The budget constraint can be expressed as: $PC = \sum_{k=1}^2 W_k L_k + \sum_{k=1}^2 \Pi_k + T$.

The aggregate consumption C is a Dixit-Stiglitz aggregator of the following form:

$$C = \left[\sum_{k=1}^2 n_k^{1/\eta} C_k^{(\eta-1)/\eta} \right]^{\eta/(\eta-1)}, \quad C_k = \left[n_k^{-1/\epsilon} \int_0^{n_k} C_{ki}^{(\epsilon-1)/\epsilon} di \right]^{\epsilon/(\epsilon-1)}.$$

where n_k is the size of sector k with the weights n_k summing up to 1. Here, η reflects the elasticity of substitution across sectors, C_k is the aggregate sectoral output, ϵ measures the elasticity of substitution within a sector, and $C_{k,i}$ denotes goods produced by firm i in sector k . In this static model, the utility-maximizing household faces two intra-temporal optimization problems. First, the household trades off aggregate consumption and labor decisions. Second, the household optimally allocates its consumption bundle within and across sectors. Solving the first problem gives rise to the labor supply equations:

$$\frac{U_{L_k}}{U_C} = -\frac{W_k}{P}, \quad \forall k \in \{1, 2\}, \quad (2.1)$$

where U_{L_k} and U_C denote the marginal utilities of labor and consumption, respectively.

The solution to the second problem yields the demand for variety i within sector k (C_{ki}) and across sectors (C_k):

$$C_{ki} = \left(\frac{P_{ki}}{P_k} \right)^{-\epsilon} \frac{1}{n_k} C_k, \quad C_k = \left(\frac{P_k}{P} \right)^{-\eta} n_k C, \quad (2.2)$$

with the aggregate price P given by:

$$P \equiv \left(\sum_{k=1}^2 n_k P_k^{1-\eta} \right)^{1/(1-\eta)}, \quad P_k \equiv \left(\int_0^1 P_{ki}^{1-\epsilon} di \right)^{1/(1-\epsilon)}. \quad (2.3)$$

Here, P_k is the sectoral price and P_{ki} denotes the price of goods i in sector k .

2.2 Firms

Firms in the model face two types of frictions: information frictions and nominal frictions. Nominal frictions are introduced to capture the infrequent price adjustments observed in the data. We assume nominal rigidities à la Calvo, where firms in sector k can reset their prices optimally with probability $1 - \theta_k$. With the remaining probability (θ_k), firms are constrained to maintain their initial prices. Although labeled as a static model, it can also be interpreted as a two-period model where the economy is in its steady state in period 0, with all firms featuring the same initial prices P^{ss} .

Firms acquire information after the resolution of pricing uncertainty. This timing assumption implies that, ex post, after the resolution of pricing uncertainty, firms that are unable to reset prices do not engage in pricing or information acquisition decisions. In contrast, price-resetting firms make two key choices: endogenous information acquisition and their optimal prices. We now discuss each of these decisions in detail.

A firm i in sector k (denoted as firm ki) produces differentiated goods (Y_{ki}) according to the following production function:

$$Y_{ki} = e^{a_k} L_{ki}, \quad (2.4)$$

where a_k represents sectoral productivity, and L_{ki} is the labor input of firm ki . Analogous to aggregate consumption goods, aggregate output is defined as $Y \equiv \left[\sum_{k=1}^2 n_k^{1/\eta} Y_k^{(\eta-1)/\eta} \right]^{\eta/(\eta-1)}$, and $Y_k \equiv \left[n_k^{-1/\epsilon} \int_{n_k} Y_{ki}^{(\epsilon-1)/\epsilon} di \right]^{\epsilon/(\epsilon-1)}$, and aggregate labor is given by $L_k \equiv \int_0^1 L_{ki} di$.

Price-resetting firms simultaneously choose their prices and the degree of attention to macroeconomic conditions (i.e., sectoral productivities) to maximize expected real profits. The real profits, denoted as div , are defined as sales net of labor costs: $div \equiv \frac{1}{P} (P_{ki} Y_{ki} - W_k L_{ki})$, where P_{ki} is the price set by firm ki , W_k is the wage rate in sector k , and Y_{ki} is the output of firm ki .

Profit-maximizing firms face a demand curve derived from the household's optimal expenditure allocation in (2.2): $Y_{k,i} = \left(\frac{P_{ki}}{P_k} \right)^{-\epsilon} \left(\frac{P_k}{P} \right)^{-\eta} Y$. Substituting the demand curve and the production function into the profit function yields:

$$div(P_{ki}, \{P_k\}, P, \{W_k\}, Y, \{a_k\}) = \frac{1}{P} \left(P_{ki} - \frac{W_k}{e^{a_k}} \right) \left(\frac{P_{ki}}{P_k} \right)^{-\epsilon} \left(\frac{P_k}{P} \right)^{-\eta} Y. \quad (2.5)$$

We now proceed to define the firm's problem. First, we derive the firm's optimal price-

setting rule with a given information set. Subsequently, we analyze the firm's optimal information acquisition problem.

Price Setting Given the information set of a firm i in sector k , a re-optimizing firm chooses an optimal price to maximize its expected profits (div):

$$\max_{P_{ki}} E [div (P_{ki}, \{P_k\}, P, \{W_k\}, Y, \{a_k\}) \mid I_{ki}], \quad (2.6)$$

where E denotes the expectation operator, I_{ki} represents the information set of firm ki , and the profit function $div (P_{ki}, \{P_k\}, P, \{W_k\}, Y, \{a_k\})$ is defined in (2.5).

The solution to this profit-maximization problem determines the optimal price P_{ki}^* :

$$P_{ki}^* = \arg \max_{P_{ki}} E [div (P_{ki}, \{P_k\}, P, \{W_k\}, Y, \{a_k\}) \mid I_{ki}]. \quad (2.7)$$

Firms are assumed to be able to change their prices only with a probability of $1 - \theta_k$. With probability θ_k , firms are stuck with their initial level of prices, P^{ss} , which is common across all firms. The aggregate price in sector k is then given by:

$$P_k = \left[(1 - \theta_k) \int_0^1 (P_{ki}^*)^{1-\epsilon} di + \theta_k (P^{ss})^{1-\epsilon} \right]^{1/(1-\epsilon)}. \quad (2.8)$$

Endogenous Information Acquisition Firms' information acquisition choices include both the selection of a signal structure and the *amount* of information. Although these decisions are made simultaneously, we present them separately for clarity. We begin with the choice of signal structure, assuming a given amount of information.

Firms do not observe a_1 and a_2 . Instead, a firm ki has access to a pool of information about unobserved states. Before delving into firms' decisions on how much attention to allocate to a particular signal, it is crucial to first discuss *which* signals are relevant for firms.

Consider a general signal structure:

$$\mathbf{s}_{ki}^\diamond = \mathbf{M}_{ki}^\diamond \mathbf{a} + \mathbf{e}_{ki}^\diamond, \quad (2.9)$$

where $\mathbf{a} = (a_1, a_2)'$ is the vector of sectoral shocks with a variance-covariance matrix Σ_{aa} , $\mathbf{s}_{ki}^\diamond = (s_{ki,1}, s_{ki,2})'$ is the vector of signals, and $\mathbf{M}_{ki}^\diamond$ is a 2×2 matrix representing the weights

of these shocks. The vector of observational errors, $e_{ki}^\diamond = (e_{1,ki}^\diamond, e_{2,ki}^\diamond)'$, is a two-dimensional Gaussian vector with a variance-covariance matrix $\Sigma_{e,ki}^\diamond$.

The information structure is general: firms have access to a continuum of signals—the information pool—characterized by the functional form (2.9) for any $M_{ki}^\diamond$ and $\Sigma_{e,ki}^\diamond$. Both $M_{ki}^\diamond$ and $\Sigma_{e,ki}^\diamond$ are determined by firms.

Formally, the optimal signal structure choice for firms is:

$$\{M_{ki}^*, \Sigma_{e,ki}^*\} = \arg \max_{\{M_{ki}^\diamond, \Sigma_{e,ki}^\diamond\}} E [\text{div} (P_{ki}^*, \{P_k\}, P, \{W_k\}, Y, \{a_k\}) \mid I_{ki}], \quad (2.10)$$

for any given amount of information, which will be described shortly, and $I_{ki} = s_{ki}^\diamond$ is defined according to (2.9).

Equivalently, for a given amount of information, firms optimally choose to pay attention to signals of the following structure:

$$s_{ki}^* = M_{ki}^* a + e_{ki}^*, \quad (2.11)$$

where M_{ki}^* and the variance-covariance matrix of e_{ki}^* , $\Sigma_{e,ki}^*$, are determined by (2.10). Note that, in equilibrium, firms within the same sector are symmetric and therefore make identical decisions. Specifically, firms in a given sector k choose the same M_{ki}^* and $\Sigma_{e,ki}^*$.

Next, we describe firms' decisions regarding the optimal amount of information acquisition. This involves balancing the benefit of acquiring more information, which mitigates profit losses from information frictions, against the associated costs.

Conditional on resetting prices, a firm ki 's optimal amount of information acquisition solves the following problem:

$$\max_{\kappa_{ki}} E [\text{div} (P_{ki}^*, \{P_k\}, P, W_k, Y, \{a_k\}) \mid I_{ki}] - \lambda f(\kappa_{ki}), \quad (2.12)$$

given the choice of information set $I_{ki} = s_{ki}^*$, defined according to (2.11), the optimal price-setting rule (2.7), and a constraint on information flow $I(\Sigma_{aa}, \hat{\Sigma}_{aa,ki})$:

$$I(\Sigma_{aa}, \hat{\Sigma}_{aa,ki}) = \kappa_{ki}, \quad (2.13)$$

where $\hat{\Sigma}_{aa,ki}$ is the posterior variance-covariance matrix with $\hat{\Sigma}_{aa,ki}^{-1} = \Sigma_{aa}^{-1} + (M_{ki}^*)' \Sigma_{e,ki}^{*-1} M_{ki}^*$. Here, κ_{ki} is the control variable that measures the *amount of information* a price-resetting firm i in sector k acquires, $f(\kappa_{ki})$ denotes a cost function, and λ is a key parameter that

characterizes the cost of information acquisition, thus determining the degree of information frictions in the economy.

According to information theory:

$$I(\Sigma_{aa}, \hat{\Sigma}_{aa,ki}) = \frac{1}{2} \log_2 \left(\frac{|\Sigma_{aa}|}{|\hat{\Sigma}_{aa,ki}|} \right), \quad (2.14)$$

where $|\cdot|$ denotes the determinant operator. Intuitively, $I(\Sigma_{aa}, \hat{\Sigma}_{aa,ki})$ measures the amount of information contained in s_{ki}^* as the difference between unconditional uncertainty (measured by entropy) and conditional uncertainty.

The solution to this optimization problem determines κ_{ki}^* :

$$\kappa_{ki}^* = \arg \max_{\kappa_{ki}} E [\text{div} (P_{ki}^*, \{P_k\}, P, W_k, Y, \{a_k\}) \mid I_{ki}] - \lambda f(\kappa_{ki}). \quad (2.15)$$

The paper emphasizes firms' choices of κ_{ki} , or equivalently, the *values* of $\Sigma_{e,ki}^*$ and M_{ki}^* .

2.3 Linear-Quadratic Approximation of the Economy

To solve the model analytically, this section log-linearizes agents' optimality conditions. To this end, we make use of the second-order approximation of firms' profit function to solve for the firms' information acquisition problem, and the second-order approximation of the welfare function to solve for the central bank's optimal monetary policy.

Appendix B.2 presents the linearized household optimality conditions and firms' price-setting equations, which are standard. For example, firms' optimal price setting is characterized by:

$$p_{ki}^* = E[p + x + u_k \mid I_{ki}] \quad (2.16)$$

where $p = n_1 p_1 + n_2 p_2$ is the aggregate price, $x = y - y^N$ is the output gap, and $y^N \equiv n_1 a_1 + n_2 a_2$ is the natural level of output, defined as the level of output that would prevail in the absence of nominal and information frictions. As explained above, we consider the case where the sizes of the two sectors are the same, i.e., $n_1 = n_2 = 0.5$, to exclude heterogeneity arising from sector sizes and to focus our analysis on heterogeneous nominal frictions. As a result, $u_1 = -u_2 = u$, with $u = 0.5(a_2 - a_1)$ capturing the asymmetric nature of sectoral shocks. Under a price stabilization policy, u is sufficient to solve the model (see, e.g., [Ou et al. 2021](#)). From this point onward, we focus on the relevant un-

known state, u , which is assumed to be drawn from a normal distribution: $u \sim N(0, \sigma_u^2)$.

Below, we focus on the novel aspect of our paper: firms' linearized optimal information choices.

Firms' Optimal Choice of Signal Structure Lemma D.1 in the Appendix D shows that, given the general signal structure (2.9), a firm ki optimally chooses to allocate its attention to one signal of the following form:

$$s_{ki}^* = u + e_{ki}, \quad e_{ki} \sim N(0, \sigma_{e,k}^2). \quad (2.17)$$

The intuition for this result is that, in our model, u is the sufficient unobserved state that determines firms' profits. Other signals or other linear combinations of a_1 and a_2 are only useful to the extent that they provide information about u . Therefore, focusing on s_{ki} is sufficient to minimize firms' profit losses. This implies that firms choose their information set to be:

$$I_{ki} = s_{ki}^*, \quad (2.18)$$

with s_{ki}^* defined according to (2.17).

Firms' Optimal Information Acquisition Following Mackowiak and Wiederholt (2009) and Mackowiak et al. (2009), we transform the profit maximization problem into a minimization problem, where the objective function is the profit loss plus the cost of paying attention. The details can be found in Appendix C.2. As a result, the optimal information acquisition problem becomes a solution to a linear-quadratic problem. Hence, the non-linear equation (2.15), which characterizes the optimal amount of information, is rewritten as:

$$\kappa_{ki}^* = \arg \min_{\kappa_{ki}} E \left[\frac{\epsilon - 1}{2} (p_{ki}^\diamond - p_{ki}^*)^2 \mid s_{ki}^* \right] + \lambda f(\kappa_{ki}). \quad (2.19)$$

subject to the constraint on information flow $I(\sigma_u^2, \hat{\sigma}_{u,ki}^2)$:

$$I(\sigma_u^2, \hat{\sigma}_{u,ki}^2) \leq \kappa_{ki}, \quad (2.20)$$

where $\hat{\sigma}_{u,ki}^2 = \text{var}(u|s_{ki}^*) = E[(u - E[u|s_{ki}^*])^2|s_{ki}^*]$. Once firms have determined their optimal attention choices (i.e., precisions of s_{ki}^*), they update their beliefs using Bayes' rule:

$$E(u | s_{ki}^*) = K_k s_{ki}^*, \quad (2.21)$$

where $K_k \equiv \frac{\sigma_u^2}{\sigma_u^2 + \sigma_{e,k}^2}$ denotes the Kalman gain. It is important to note that in firms' attention choice, selecting the constraint on information flow κ_k , posterior uncertainty $\hat{\sigma}_{u,ki}^2$, the precision of the signal $1/\sigma_{e,k}^2$, or the Kalman gain K_k are equivalent.

2.4 Central Bank: Optimal Monetary Policy

This section presents the key optimization problem of the paper: optimal monetary policy. We proceed with a description of the alternative types of monetary policy and the microfounded central bank objective function, followed by the formal definition of optimal monetary policy.

Price Stabilization Policy Throughout the paper, we consider two alternative classes of monetary policies.

The first class of monetary policy rules considered is the price index stabilization policy, which gives the central bank the capacity to stabilize a specified aggregate price index. Specifically, the policy is formulated as:

$$\omega_1 p_1 + \omega_2 p_2 = 0, \quad (2.22)$$

where the central bank's target is the price index $\omega_1 p_1 + \omega_2 p_2$, and the central bank's choice of weights $\{\omega_1, \omega_2\}$ plays a crucial role. Intuitively, the terms $\frac{\omega_1}{\omega_1 + \omega_2}$ and $\frac{\omega_2}{\omega_1 + \omega_2}$ represent the relative weights assigned to the prices in sector 1 and sector 2, respectively. Consequently, the policy rule in (2.22) can be equivalently expressed as:

$$\omega p_1 + (1 - \omega) p_2 = 0, \quad (2.23)$$

where $\omega \equiv \frac{\omega_1}{\omega_1 + \omega_2}$. From this point onward, we use ω as a sufficient statistic to characterize a price stabilization policy. If the optimal policy prescribes $\omega > 0.5$, we interpret this as $\omega_1 > \omega_2$. Throughout the paper, we use ω and $\{\omega_1, \omega_2\}$ interchangeably to refer to the central bank's choice.

Within the framework of the price index stabilization policy, we examine three distinct

policies. The first is referred to as the Consumer Price Index (CPI) stabilization policy and is represented by $\omega_1 = n_1$ and $\omega_2 = n_2$. The second is the Optimal Price Index (OPI) stabilization policy, in which $\{\omega_1, \omega_2\}$ are optimally selected to maximize the ex-ante (unconditional) expected welfare of the household $E\{U(C, \{L_k\})\}$ in the model with information and nominal frictions. The third is the Sticky Price Index (SPI) stabilization policy. According to the SPI stabilization policy, the central bank computes $\{\omega_1, \omega_2\}$ to maximize the ex-ante expected welfare in the model with nominal rigidities but without information frictions.

The second class of monetary policy considered is the optimal Ramsey policy, according to which the central bank maximizes $E\{U(C, \{L_k\})\}$ without *necessarily* restricting itself to the class of price index stabilization policies. Next, we formally define the central bank's objective function and the optimal monetary policy.

Central Bank's Objective Function We transform the central bank's welfare maximization problem into a minimization problem, where the objective function is a welfare loss function. The welfare loss function, in an economy with dispersed beliefs, is derived as the second-order approximation of the household's utility function following [Ou et al. \(2021\)](#):

$$E\mathcal{L} = E\left[\epsilon \sum_{k=1}^2 n_k \int_i (p_{ki} - p_k)^2 di + x^2 + \eta n_1 n_2 (\tilde{p}_1 - \tilde{p}_2)^2\right], \quad (2.24)$$

where $\tilde{p}_1 - \tilde{p}_2 \equiv p_1 - p_2 - (p_1^N - p_2^N)$. Here, p_k^N represents the natural level of the sectoral price, defined as the price level that would prevail in the absence of nominal and informational frictions.⁴ Here, $E\mathcal{L}$ denotes the expected welfare loss expressed as a fraction of steady-state consumption (up to additive terms independent of policy).⁵ Note that the expectation is taken unconditionally to reflect that the central bank must make a policy decision, i.e., choose the price index, before the realization of shocks and signals.

The welfare loss function consists of three components: the output gap x^2 , the cross-sector relative price gap $\eta n_1 n_2 (\tilde{p}_1 - \tilde{p}_2)^2$, and the within sector price dispersion. These components constitute the well-known welfare loss function in a multi-sector setting. See, for example, [Woodford \(2003\)](#) for a textbook treatment.

The key mechanisms highlighted in this paper operate through the price dispersion

⁴Note that $p_1^N - p_2^N = 2u$ when $n_1 = n_2 = 0.5$.

⁵The welfare loss function is obtained under the standard assumption of the presence of optimal wage subsidies that eliminate distortions arising from monopolistic competition in the steady state.

component, which can be decomposed into Calvo price dispersion and dispersed-belief price dispersion:

$$E \epsilon \underbrace{\sum_{k=1}^2 n_k \int_i (p_{ki} - p_k)^2 di}_{\text{Price dispersion}} = \underbrace{\epsilon \sum_{k=1}^2 n_k (1 - \theta_k) \theta_k E[p_k^{*2}]}_{\text{Calvo price dispersion}} + \underbrace{\epsilon \sum_{k=1}^2 n_k (1 - \theta_k) E \left(\int_i (p_{ki}^* - p_k^*)^2 di \right)}_{\text{Dispersed-belief price dispersion}}.$$

The Calvo price dispersion captures the price dispersion between price-resetting firms and sticky price firms due to nominal rigidity. The dispersed-belief price dispersion measures the price dispersion within the price-resetting firms due to the heterogeneous signals they receive.

There are two points that are worth clarifying. First, the Calvo price dispersion component can be rewritten in terms of aggregate prices (the counterpart of inflation in the static setting),⁶ and thus the welfare loss function contains the familiar terms: aggregate prices and output gaps. Second, however, dispersed beliefs give rise to an additional component that cannot be reduced to aggregate prices. Moreover, in this multi-sector setting, in contrast to a one-sector economy, the divine coincidence—the stabilization of prices being associated with the closing of the output gap—does not hold due to the presence of sector-specific shocks.⁷ Therefore, in general, sector-specific shocks create welfare losses by generating cross-sector relative price gaps, as well as output gaps and price dispersion.

Optimal Monetary policy We are now ready to formally state the optimal monetary policy problem.

Given the quadratic welfare loss function and the linearized firms' optimality condition derived in Appendix B.2, the OPI stabilization policy solves the following linear-quadratic minimization:

$$\min_{\omega} E\mathbb{L}$$

subject to: Monetary policy rule (2.23): $\omega p_1 + (1 - \omega)p_2 = 0$ and firms' optimality conditions that jointly characterize the aggregate supply side of the economy.⁸ In contrast, the

⁶Using the fact that $p_k = (1 - \theta_k)p_k^*$.

⁷This can be seen from the aggregate supply curve, Equation B.8: setting $p = x = 0$ is not possible in the presence of sector-specific shocks, i.e., when $u \neq 0$.

⁸See Appendix B.2 for details of linear constraints.

Ramsey optimal monetary policy is defined as follows:

$$\min_{\phi} E\mathbb{L}$$

subject to the same constraints, without restricting to a price index stabilization rule i.e., without the constraint (2.23). Instead, the central bank optimally targets aggregate nominal demand, given by $p + x = \phi u$, by choosing the ϕ to minimize the unconditional welfare loss. Notably, under the optimal monetary policy, the central bank determines aggregate nominal demand as an arbitrary function of the sufficient state in an unrestricted manner.

2.5 Timing of the Model

This static model consists of five stages. In the first stage, the central bank chooses a monetary policy rule. In the second stage, firms choose a signal structure and make a state-contingent information acquisition plan. Specifically, firms make an advance plan regarding the amount of information, κ_{ki} , to acquire at a subsequent stage—stage four—contingent on whether they are able to reset prices during that stage. In stage three, nature draws unobserved fundamental shocks. In the same stage, the Calvo uncertainty—whether firms can reset prices or not—is resolved. In stage four, price-resetting firms acquire information according to their plans made in stage two. The cost of information acquisition $\lambda f(\kappa)$ is paid at this stage. Firms then make a price-setting plan (i.e., goods-level supply curve) conditional on their information set. Finally, in stage five, the representative household observes the state of the economy and makes consumption and labor decisions, while the goods and labor markets clear.

Appendix B.3 formally defines the equilibrium of the model, which is consistent with the optimization of the household, firms, and the central bank.

3 Results

3.1 The Key Mechanism

Before delving into the mechanism, it is useful to describe a few relevant policy functions (solutions) of the model.

Policy Functions As shown in Appendix C.3, the solution of the model features the following optimal price-setting policy functions:

$$p_{1i}^* = \varphi_1 E(u \mid s_{1i}), \quad p_{2i}^* = -\varphi_2 E(u \mid s_{2i}), \quad (3.1)$$

where firms update their beliefs according to (2.21). Here, φ_1 and φ_2 characterize firms' responsiveness to the *perceived* state of the economy. While the details of these policy functions are provided in the Appendix C.3, we highlight that φ_1 and φ_2 are functions of the underlying nominal rigidities in the economy, θ_1 and θ_2 .

With the policy functions and the information flow constraint (2.20), the optimal amount of information (2.19) can be rewritten as:

$$\kappa_{ki}^* = \arg \min_{\kappa_{ki}} \frac{(\epsilon - 1)}{2} \varphi_k^2 2^{-2\kappa_{ki}} \sigma_u^2 + \lambda f(\kappa_{ki}). \quad (3.2)$$

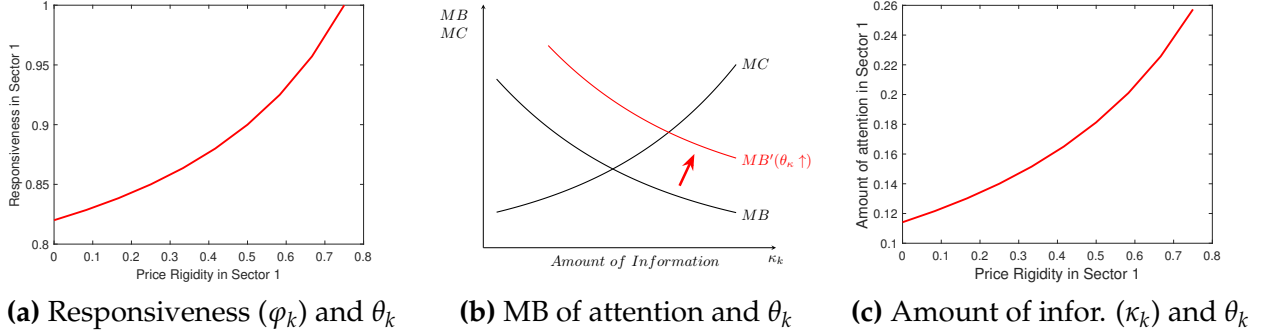
Combining (3.1) with (3.2), it becomes evident that nominal rigidities θ_k affect information acquisition decisions κ_{ki}^* through their impact on the pricing policy functions φ_k . Proposition 1 below formalizes this result. Note that, given the model parameters, monetary policy may lead the economy into a region where firms acquire no information and the output gap is indeterminate. This occurs whenever ω falls outside of $(0, 1)$: see Appendix D.9 for details. Throughout the paper, we focus on the optimal monetary policy that avoids such indeterminacy; i.e., we restrict our attention to the OPI with $\omega \in (0, 1)$. Moreover, given the property of the symmetric equilibrium within a sector, the attention choice within price-resetting firms in each sector is homogeneous, such that $\kappa_{ki}^* = \kappa_k^*$. Henceforth, we drop the i subscript to ease the notational burden.

The Key Mechanism: Information Acquisition and Nominal Rigidity The endogenous relationship between nominal rigidities and attention choice is central to the results presented in the next section. The following proposition summarizes the key mechanism:

Proposition 1. *The responsiveness of prices to expected shocks (φ_k) and the amount of information (κ_k) acquired by firms in sector k increase with θ_k , for any given monetary policy rule $\omega \in (0, 1)$, if and only if $\frac{f''(\kappa)}{f'(\kappa)} > -2 \ln 2$. That is, $\frac{\partial \varphi_k}{\partial \theta_k} > 0$, $\frac{\partial \kappa_k}{\partial \theta_k} > 0$, and $\frac{\partial \sigma_{e,k}}{\partial \theta_k} < 0$, if and only if $\frac{f''(\kappa)}{f'(\kappa)} > -2 \ln 2$.*

Proof. See Appendix D.2. □

Figure 1: Illustration of Proposition 1



Note: This figure illustrates how nominal rigidities (θ_1) affect information choice (κ_1 and $\hat{\sigma}_{u,1}$). Panel (b) provides a simple example illustrating how θ_1 affects the marginal benefit (MB) of acquiring information in partial equilibrium. Panels (a) and (c) plot the solutions of a calibrated model. Specifically, Panel (a) and Panel (c) plot the responsiveness φ_1 and the information choice κ_1 as a function of θ_1 , respectively. The model is calibrated as follows: $\theta_2 = 0.75, \epsilon = 11, \eta = 1, \sigma_u^2 = 0.02, n_1 = n_2 = 0.5$. Moreover, we assume $f(\kappa_k) = 2^{2\kappa_k}$, with the parameter determining the attention cost (λ) set to 0.049, aligning K^s with [Coibion and Gorodnichenko \(2012\)](#)'s empirical estimate of the Kalman gain (0.3).

The main takeaway is that firms operating in sectors with *stickier prices* tend to pay *more attention* when they are able to reset prices.

To understand how nominal rigidities affect attention choice, note that in a sector with stickier prices, re-optimizing firms respond more strongly to the perceived state of the economy, i.e., $\frac{\partial \varphi_1}{\partial \theta_1} > 0$. This result is confirmed by Figure 1a, which plots the solution for φ_1 as a function of θ_1 . The intuition is as follows: consider a change in u driven by a positive productivity shock in Sector 1, which affects the marginal cost of production through two channels. First, a positive productivity (supply) shock directly reduces marginal cost and therefore leads firms to *cut* prices. Second, increased productivity raises aggregate demand and, in turn, the nominal wage, which is part of marginal cost.⁹ In response to higher wages resulting from stronger demand, price-resetting firms in Sector 1 have an incentive to *increase* their prices. The second channel is a general-equilibrium (GE) force operating through aggregate demand, which offsets the first, direct effect. When nominal rigidity is high, the aggregate price in sector 1 declines less in response to a positive productivity shock, dampening the demand expansion from lower prices. Consequently, the nominal wage rises less, thereby dampening the offsetting GE force. In a summary, the strength of this GE force depends on the degree of nominal rigidity. Ultimately, with more

⁹The linearized marginal cost of production equals the nominal wage minus the marginal product of labor.

rigid prices, price-resetting firms in Sector 1 reduce prices by more: $\frac{\partial \varphi_1}{\partial \theta_1} > 0$. A similar intuition holds when considering changes in u driven by a productivity shock in Sector 2.¹⁰ Note that these channels are also present in a textbook one-sector New Keynesian model with full information. Appendix D.2 shows that the same mechanisms hold in our multi-sector New Keynesian model with endogenous information acquisition.

The more sensitive firms' prices are to the perceived state of the economy, the greater the marginal benefit (MB) of acquiring information. In fact, for firms in sector 1, the marginal benefit of acquiring information is proportional to φ_1 . Intuitively, in the limiting case where $\varphi_1 = 0$, firms have no incentive to adjust prices, and the marginal benefit of acquiring information becomes zero.

The above results establish a positive relationship between an individual firm ki 's marginal benefit of acquiring information and the degree of nominal rigidity in sector k . This is shown graphically in Figure 1b, where the MB curve shifts upward as nominal rigidities increase ($\theta_k \uparrow$). By contrast, the marginal cost (MC) curve is not affected by nominal rigidities. Therefore, in equilibrium, firms operating in a sector with stickier prices pay more attention when they are allowed to adjust prices. This is evidenced by $\frac{\partial \kappa_1}{\partial \theta_1} > 0$ and $\frac{\partial \hat{\sigma}_{u,1}}{\partial \theta_1} < 0$, as shown in Figure 1c.

Note that the mechanism described in Proposition 1 relies on a condition for the cost function $f(\kappa)$: $\frac{f''(\kappa)}{f'(\kappa)} > -2 \ln 2$. First, two common cost functions employed in the literature, an exponential form $f(\kappa_k) = 2^{2\kappa_k}$ proposed by Paciello and Wiederholt (2014), and a simple linear cost function $f(\kappa_k) = \kappa_k$, both satisfy this condition.

Second, Proposition 1 also holds under a more intuitive but more restrictive condition: $f' > 0$ and $f'' \geq 0$. This condition implies that the marginal cost of information acquisition is positive and non-declining. The intuition is illustrated in Figure 1b: as long as the marginal cost of information acquisition is positive and non-declining, an upward shift in the marginal benefit curve due to a higher θ necessarily increases the amount of information acquisition, regardless of the specific functional form of the marginal cost curve. Note that the two cost functions mentioned also satisfy this more restrictive condition.

In the baseline model, following Paciello and Wiederholt (2014), we assume an exponential cost function $f(\kappa_k) = 2^{2\kappa_k}$ due to its analytical convenience, which is important

¹⁰When the increase in u is driven by a productivity shock in Sector 2, the supply-side effect is absent for price-resetting firms in Sector 1. In this case, the rise in nominal wages—induced by stronger aggregate demand—provides an incentive for firms in Sector 1 to raise prices, thereby partially offsetting the initial demand expansion. When nominal rigidity is high, prices adjust less, and this offsetting channel becomes weaker. As a result, price-resetting firms set a higher price.

for deriving the main proposition presented below. Numerically, our findings are robust to alternative cost functions.

Given the assumed cost function, the optimal attention choice corresponds to the following level of attention, summarized by the Kalman gain: $K_k = \max\{1 - (\frac{2\lambda}{(\epsilon-1)\phi_k^2\sigma_u^2})^{\frac{1}{2}}, 0\}$. Note that the Kalman gain K_k must be non-negative.

To understand the parameters that determine the degree of information frictions, it is useful to introduce the following definition:

$$K^s \equiv 1 - \left(\frac{2\lambda}{(\epsilon-1)\sigma_u^2} \right)^{\frac{1}{2}}, \quad (3.3)$$

where K^s represents the Kalman gain selected by firms in a counterfactual symmetric economy with $\theta_1 = \theta_2$ and $\omega = 0.5$. The analytical expression in (3.3) indicates that the degree of information frictions is proportional to the cost of information acquisition: K^s decreases as λ increases.

3.2 An Equivalence Result: OPI and Ramsey Optimal Policy

We now establish the equivalence between the OPI stabilization policy and the Ramsey optimal monetary policy.

Lemma 3.1. *A simple price stabilization rule: $\omega_1^*p_1 + \omega_2^*p_2 = 0$ can replicate the Ramsey optimal monetary policy.*

Proof. See Appendix D.3 □

The details of the proof are provided in D.3. A similar result was shown by La'O and Tahbaz-Salehi (2022) in a model with exogenous information and by Fang et al. (2024) in a setting with endogenous information acquisition and flexible prices.

The intuition for why an OPI stabilization policy can replicate the Ramsey optimal policy is as follows. For a given set of policy coefficients $\{\omega_1, \omega_2\}$, an equilibrium is determined. More importantly, the converse is also true. In other words, the set of equilibria is spanned by the possible values of $\{\omega_1, \omega_2\}$, including the Ramsey optimal equilibrium. Therefore, for any equilibrium implemented under the Ramsey policy, it is always possible to find a set of ω_1^* and ω_2^* that achieves the same equilibrium allocation.

In light of this lemma, we use the terms OPI stabilization policy and optimal monetary policy interchangeably to describe the Ramsey optimal monetary policy.

Table 1: Optimal Monetary Policy under Different Assumptions

	(1) Full Information	(2) Exogenous Infor- mation	Infor- mation	(3) Endogenous Infor- mation (Baseline)	(4) Endogenous Infor- mation (Dynamic)
ω_1^* vs. ω_2^* :	<	<		>	>
Condition:	Unconditionally	Unconditionally		Sufficiently Costly Information Acqui- sition	Standard Calibra- tion

3.3 The Main Results: Optimal Monetary Policy

At this point, it is helpful to formally state the research question. Consider an economy with two sectors that are otherwise symmetric except for differing degrees of nominal rigidity, i.e., $\theta_1 \neq \theta_2$, and without loss of generality, we consider the case $\theta_1 < \theta_2$. The research question is: Is $\omega_1^* < \omega_2^*$ when $\theta_1 < \theta_2$?

We now present the main findings.

Proposition 2 (Conventional View). *Consider an economy with **exogenous** information frictions where nominal rigidities are the only source of heterogeneity across sectors. The optimal monetary policy attaches a larger weight to the sector with stickier prices: i.e., $\omega_1^* < \omega_2^*$ if $\theta_1 < \theta_2$.*

Proof. See Appendix D.4 □

Proposition 2 re-establishes the conventional view in our setting with exogenous information frictions, where $\sigma_{e,1} = \sigma_{e,2}$ are set to exogenous values. Note that a model with full information is nested as a special case where $\sigma_{e,1} = \sigma_{e,2} = 0$. Columns (1) and (2) in Table 1 summarize these findings: with full information or exogenous information frictions, the optimal monetary policy attaches a larger weight to the sector with stickier prices, i.e., $\omega_1^* < \omega_2^*$ if $\theta_1 < \theta_2$, *unconditionally*.

In contrast, when information acquisition is an endogenous choice, and $\sigma_{e,1}$ and $\sigma_{e,2}$ are endogenous, the policy prescription changes dramatically.

Proposition 3 (Main Result). *Consider the economy with endogenous information acquisition*

and $\theta_1 < \theta_2$. There exists a threshold value $\bar{K}$ such that

$$\omega_1^* > \omega_2^* \quad \text{if and only if} \quad K^s < \bar{K}.$$

That is, the optimal monetary policy attaches a smaller weight to the sector with stickier prices if and only if information acquisition is sufficiently costly.

Proof. See Appendix D.5 □

Proposition 3 shows that, with endogenous information acquisition, in an economy where $\theta_1 < \theta_2$, $\omega_1^* > \omega_2^*$ if and only if information acquisition is sufficiently costly. This result is presented in Column (3) of Table 1. Note that K^s , defined in (3.3), is a measure of the degree of information friction or the cost of information acquisition.

Understanding the Results via a Decomposition To better understand the results, we perform the following decomposition exercise, as in La'O and Tahbaz-Salehi (2022).

Recall that the total welfare loss ($\mathbb{L}$) function (2.24) is composed of the dispersed-belief price dispersion ($\mathbb{L}_{DispB}$), the Calvo price dispersion ($\mathbb{L}_{Calvo}$), the output gap ($\mathbb{L}_{OutputGap}$), and the cross-sector relative price gap ($\mathbb{L}_{RPGap}$) components:

$$E\mathbb{L} = E\mathbb{L}_{DispB} + E\mathbb{L}_{Calvo} + E\mathbb{L}_{OutputGap} + E\mathbb{L}_{RPGap}. \quad (3.4)$$

The optimal monetary policy $\{\omega_1^*, \omega_2^*\}$ described in Propositions 2 and 3 minimizes the total welfare loss (3.4) subject to the constraints summarized in Section 2.4.

In the decomposition exercise, we compute $\{\omega_{1,DispB}^*, \omega_{2,DispB}^*\}$, $\{\omega_{1,Calvo}^*, \omega_{2,Calvo}^*\}$, $\{\omega_{1,OutputGap}^*, \omega_{2,OutputGap}^*\}$, and $\{\omega_{1,RPGap}^*, \omega_{2,RPGap}^*\}$ separately, by minimizing each corresponding *sub-component* of the total welfare loss function subject to the same constraints. Formally:

$$\{\omega_{1,Sub}^*, \omega_{2,Sub}^*\} = \arg \min_{\{\omega_{1,Sub}, \omega_{2,Sub}\}} E\mathbb{L}_{Sub}, \quad (3.5)$$

subject to the constraints summarized in Section 2.4 for each sub-component $Sub \in \{DispB, Calvo, OutputGap, RPGap\}$.

It is straightforward to show that:

$$\omega_k^* = \omega_{k,DispB}^* + \omega_{k,Calvo}^* + \omega_{k,OutputGap}^* + \omega_{k,RPGap}^*. \quad (3.6)$$

That is, the optimal monetary policy $\{\omega_k^*\}$ can be decomposed into components arising from each part of the welfare loss function. This decomposition strategy leverages the additive property of the total welfare loss function illustrated by (3.4).

Lemma 3.2 presents the decomposition results.

Lemma 3.2. *Consider an economy where nominal rigidities are the only source of ex-ante heterogeneity across sectors and $\theta_1 < \theta_2$. When information frictions are exogenous:*

- $\omega_{1,Calvo}^* < \omega_{2,Calvo}^*, \omega_{1,OutputGap}^* < \omega_{2,OutputGap}^*, \text{ and } \omega_{1,RPGap}^* < \omega_{2,RPGap}^*;$
- $\omega_{1,DispB}^* = \omega_{2,DispB}^*.$

In contrast, with endogenous information acquisition:

- $\omega_{1,Calvo}^* < \omega_{2,Calvo}^*, \omega_{1,OutputGap}^* < \omega_{2,OutputGap}^*, \text{ and } \omega_{1,RPGap}^* < \omega_{2,RPGap}^*;$
- $\omega_{1,DispB}^* > \omega_{2,DispB}^*.$

Moreover, the difference between $\omega_{1,DispB}^*$ and $\omega_{2,DispB}^*$ is amplified when K^S is smaller.

Proof. See Appendix D.6 □

First, consider the conventional channels that operate through $E\mathbb{L}_{Calvo}$, $E\mathbb{L}_{OutputGap}$, and $E\mathbb{L}_{RPGap}$. Lemma 3.2 shows that, when $\theta_1 < \theta_2$, we have $\omega_{1,Calvo}^* < \omega_{2,Calvo}^*, \omega_{1,OutputGap}^* < \omega_{2,OutputGap}^*, \text{ and } \omega_{1,RPGap}^* < \omega_{2,RPGap}^*$ in both the exogenous and endogenous information settings. In other words, according to these three well-studied components of the welfare loss function, optimal monetary policy tends to assign a larger weight to the sector with stickier prices. Moreover, endogenous information does not alter this property. This result is unsurprising, as nominal rigidity is the main friction driving the aggregate gaps: Calvo price dispersion, output gap, and cross-sector relative price gap. According to these channels, price fluctuations are more costly in sectors with more rigid prices. Consequently, optimal monetary policy counteracts price fluctuations more strongly in these sectors.

Second, Lemma 3.2 shows that the key component driving the discrepancy between Propositions 2 and 3 is the dispersed-belief welfare loss component, which arises in models with information frictions and dispersed beliefs. In a model with exogenous information frictions, where the only source of heterogeneity is nominal rigidities, the $\mathbb{L}_{DispB}$ component tends to assign equal weight to price fluctuations in both sectors, i.e., $\omega_{1,DispB}^* =$

$\omega_{2,DispB}^*$. This component, combined with the other components, establishes the unconditional result in Proposition 2.

In stark contrast, $\omega_{1,DispB}^* > \omega_{2,DispB}^*$ when information acquisition is endogenous. The intuition is that heterogeneity in nominal frictions ($\theta_1 < \theta_2$) endogenously generates heterogeneity in information frictions. Specifically, as shown in Proposition 1, firms in a more flexible price sector endogenously acquire less information. As a result, there is greater disagreement among price-resetting firms, which ultimately increases the dispersed-belief price dispersion component. This channel explains why price fluctuations in more flexible price sectors impose greater welfare costs due to disagreements in optimal prices among price-resetting firms. Consequently, optimal policy assigns a larger weight to Sector 1's prices.

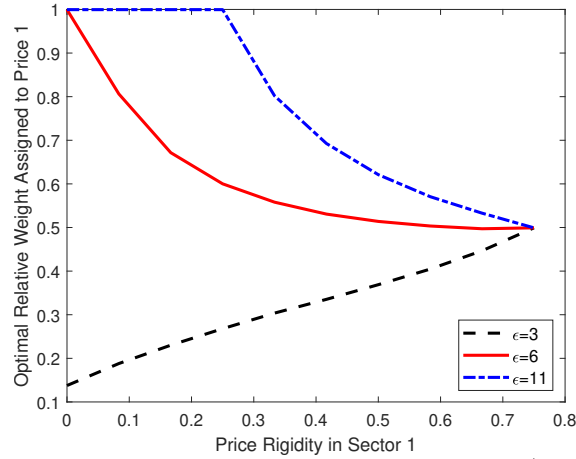
Naturally, the dispersed-belief component becomes more important when information acquisition is more costly. Formally, Lemma 3.2 shows that the difference between $\omega_{1,DispB}^*$ and $\omega_{2,DispB}^*$ is amplified when K^S is smaller. This explains the condition in Proposition 3, where the channel we highlight offsets other channels when $K^S < \bar{K}$.

Whether the condition $K^S < \bar{K}$ is satisfied in reality is a quantitative question, which we address below. Qualitatively, the larger $\bar{K}$, the more likely the condition is satisfied. Since $\frac{\partial \bar{K}}{\partial \epsilon} > 0$, a higher ϵ increases the importance of the channel we highlight. This is because our channel operates through the price dispersion component of the welfare loss, whose relative importance for total social welfare rises with ϵ . With this background, we evaluate the optimal policy under alternative values of ϵ in our quantitative exploration.

A Quantitative Exploration We calibrate the model to assess the quantitative relevance of the channel we propose. Specifically, we set $\theta_2 = 0.75$, which corresponds to an average price duration of one year in the sticky price sector. Following [Hobijn and Nechio \(2019\)](#)'s empirical estimate, we set the cross-sector elasticity of substitution η to 1. The parameter σ_u^2 is irrelevant for the qualitative findings of the paper, and we set it to 0.02. Moreover, the parameter that determines the attention cost (λ) is calibrated to align K^S with [Coibion and Gorodnichenko \(2012\)](#)'s empirical estimate of the Kalman gain (0.3).

Figure 2 illustrates the optimal monetary policy for the calibrated model. Specifically, it plots the optimal monetary policy summarized by $\frac{\omega_1^*}{\omega_1^* + \omega_2^*}$ as a function of price rigidity in sector 1 (θ_1) under three alternative calibrations of ϵ . The figure can be read from the right end of the horizontal axis to the left. The point where the three curves coincide corresponds to the scenario where $\theta_1 = \theta_2 = 0.75$. In this scenario, sectors are symmetric

Figure 2: Optimal Monetary Policy



Note: This figure plots the optimal monetary policy summarized by $\frac{\omega_1^*}{\omega_1^* + \omega_2^*}$ as a function of price rigidity in sector 1 (θ_1) under three alternative calibrations of ϵ . The calibration of the remaining parameters of the model is as follows: $\theta_2 = 0.75$, $\eta = 1$, $\sigma_u^2 = 0.02$, $n_1 = n_2 = 0.5$. Moreover, the parameter that determines the attention cost (λ) is calibrated to align K^s with Coibion and Gorodnichenko (2012)'s empirical estimate of the Kalman gain (0.3).

in all characteristics except that they are subject to asymmetric sectoral shocks. Consequently, the optimal policy assigns equal weights ($\omega_1^* = \omega_2^* = 0.5$) to both sectors across the three calibrations of ϵ considered. As θ_1 decreases (moving from the right end of the horizontal axis to the left), since θ_2 is held fixed, the nominal frictions between the two sectors become more heterogeneous, leading to an optimal price index that deviates from the CPI.

Importantly, the policy prescriptions depend on the calibration of ϵ , which determines the importance of the channel we emphasize. In the first scenario, where the within-sector elasticity of substitution $\epsilon = 11$, reflecting an average price markup of 10% (consistent with recent empirical evidence provided by Edmond et al. (2018)), the calibrated model features $K^s < \bar{K}$. As a result, the optimal policy assigns a larger weight $\frac{\omega_1^*}{\omega_1^* + \omega_2^*} > 0.5$ to the sector with more flexible prices, as shown by the blue line. A similar result holds when $\epsilon = 6$, another common calibration value chosen in the literature (see, e.g., Galí 2015), which corresponds to a markup of 20%. In contrast, when ϵ is sufficiently small, the $K^s < \bar{K}$ condition is violated. As a result, the policy prescription in our model aligns with the conventional view, as evident from the black dashed line.

3.4 Is the Stabilization of CPI a Good Policy?

Corollary 4 is directly implied by Proposition 3: In a special case where $K^s = \bar{K}$, the dispersed-belief component coincidentally offsets the other components. We label $K^s = \bar{K}$ as the Divine Coincidence Condition. Under this condition, the optimal monetary policy *coincides* with the CPI stabilization policy, independent of heterogeneities in nominal rigidities.

Corollary 4 (Divine Coincidence Condition). *Consider the economy with endogenous information acquisition, $\forall \theta_1, \theta_2$:*

$$\omega_1^* = \omega_2^* \quad \text{if and only if} \quad K^s = \bar{K}.$$

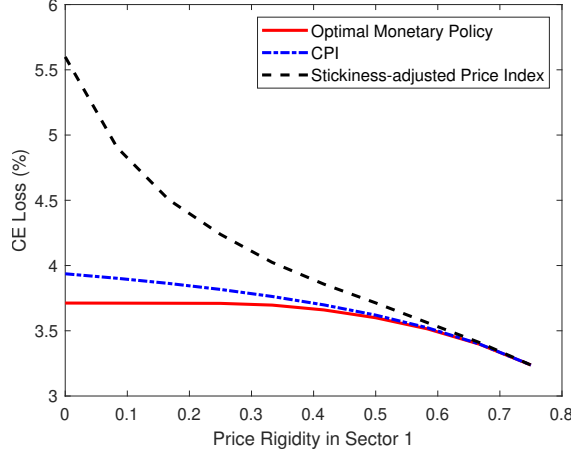
That is, the optimal monetary policy coincides with the CPI stabilization policy, independent of heterogeneities in nominal rigidities, when $K^s = \bar{K}$.

Is the Stabilization of CPI a Good Policy? The answer to this question depends on how far reality deviates from the Divine Coincidence Condition. In the next quantitative exploration, we assess the performance of CPI stabilization by comparing welfare losses under CPI stabilization with alternative policies. The calibration of the model is the same as before, and we set $\epsilon = 11$.

Figure 3 presents the results. It shows the welfare loss in the baseline model with rational inattention under three alternative monetary policy rules: the Optimal Price Index (OPI), Consumer Price Index (CPI), and Stickiness-Adjusted Price Index (SPI) stabilization policies. The OPI accounts for both endogenous information frictions and heterogeneous nominal rigidities, whereas the SPI is computed ignoring the information frictions feature of the model. The figure provides insights into how different monetary policy rules perform in terms of social welfare loss as price rigidity in sector 1 varies. Unsurprisingly, the OPI delivers the lowest welfare loss compared to the other two price index stabilization policies.

Interestingly, Figure 3 shows that CPI stabilization performs better than SPI stabilization in terms of social welfare and is not far from the welfare loss implied by the optimal policy.

Figure 3: Welfare Loss under Alternative Monetary Policies



Note: The figure shows the total social welfare loss as a function of price rigidity in sector 1 in the base-line model with rational inattention under three alternative monetary policy rules: the optimal monetary policy/OPI stabilization (red line), CPI (dotted blue line), the stickiness-adjusted price index (dashed black line). The model calibration is as follows: $\theta_2 = 0.75$, $\epsilon = 11$, $\eta = 1$, $\sigma_u^2 = 0.02$, $n_1 = n_2 = 0.5$. Moreover, the parameter determining the attention cost (λ) is calibrated to align K^s with [Coibion and Gorodnichenko \(2012\)](#)'s empirical estimate of the Kalman gain (0.3).

3.5 Monetary Policy Shapes Endogenous Attention Choice

The analysis thus far has focused on the impact of endogenous attention choice on monetary policy. In this section, we explore the reciprocal relationship between monetary policy and attention choice, specifically how monetary policy influences firms' attention choice. We also provide insights into firms' attention choices under counterfactual monetary policy scenarios.

Proposition 5. *The responsiveness of prices to the perceived state (φ_1) and the amount of information (κ_1) acquired by firms in sector 1 decrease with ω_1 . That is, $\frac{\partial \varphi_1}{\partial \omega} < 0$, $\frac{\partial \kappa_1}{\partial \omega} < 0$, and $\frac{\partial \sigma_{e,1}}{\partial \omega} > 0$.*

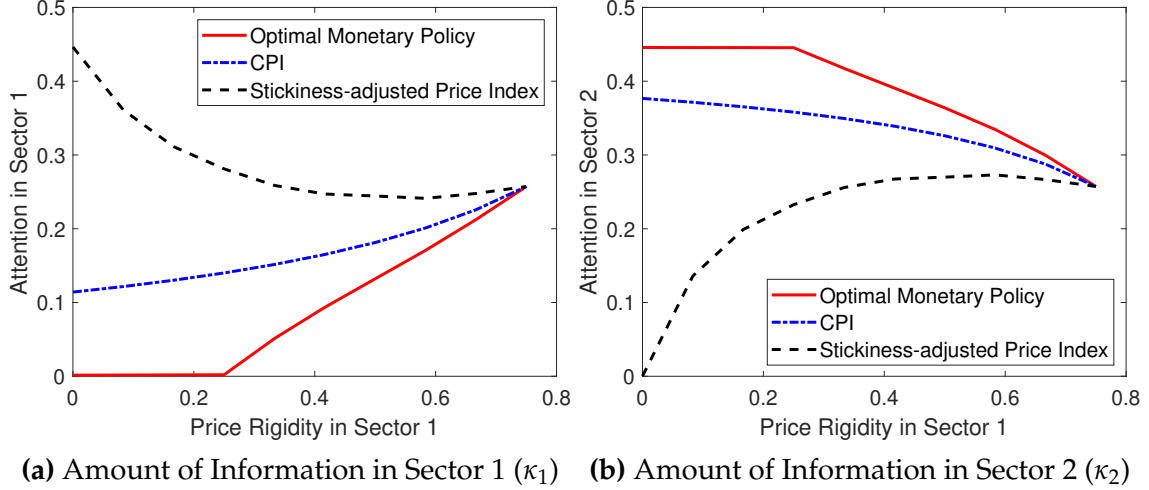
Proof. See Appendix [D.8](#). □

Proposition 5 shows that firms in a sector pay less attention to economic conditions if the central bank assigns a higher weight to the stabilization of prices in this sector.

The intuition for Proposition 5 is straightforward. First, monetary policy affects firms' responsiveness to the perceived state (φ_1). Specifically, if the central bank assigns more weight to the stabilization of sector 1's prices (ω increases), firms in this sector respond less to shocks (φ_1 decreases). Since φ_1 captures the marginal benefit of paying attention,

less responsive firms are then less attentive to economic conditions. Consequently, $\frac{\partial K_1}{\partial \omega} < 0$. As a result, $\sigma_{e,1}$ increases with ω_1 .

Figure 4: Attention Choice Under Alternative Policy Rules



Note: This figure plots the amount of information (κ_{ki}) as a function of price rigidity in sector 1 in the baseline model with rational inattention under three alternative monetary policy rules: the stabilization of the optimal price index (red line), CPI (dotted blue line) and stickiness-adjusted price index (dashed black line). The calibration of the model is as follows: $\theta_2 = 0.75$, $\epsilon = 11$, $\eta = 1$, $\sigma_u^2 = 0.02$, $n_1 = n_2 = 0.5$. Moreover, the parameter determining the attention cost (λ) is calibrated to align K^s with [Coibion and Gorodnichenko \(2012\)](#)'s empirical estimate of the Kalman gain (0.3).

This result suggests a potential risk of adopting the optimal price index recommended by the baseline model: firms in a sector with more flexible prices may pay very little attention to overall economic conditions. Figure 4 illustrates the amount of information firms acquire (κ_{ki}) as a function of price rigidity in sector 1 in the rational inattention model under three monetary policy rules: the optimal monetary policy/OPI stabilization (red line), CPI (dotted blue line) and the stickiness-adjusted price index (dashed black line). The left and right panels show the attention choices for firms in Sectors 1 and 2, respectively.

Under CPI stabilization, firms in a sector with more flexible prices pay less attention to macroeconomic conditions. The optimal monetary policy leads to even less information acquisition by firms in the more flexible price sector compared to the CPI stabilization policy. In contrast, under the SPI stabilization policy, firms in the stickier price sector may acquire very little information. Compared to the extreme attention choice outcomes under the OPI and SPI stabilization policies, firms in the CPI stabilization world make more balanced attention choices.

3.6 Discussions: Alternative Timing

Throughout the paper, we assume the timing convention that information is obtained upon the need to change prices. An alternative timing assumption is that information is acquired before a firm knows whether it can adjust its price.

Our preferred timing assumption is justified for three reasons. First, if firms were allowed to choose *the timing* of information acquisition, they would optimally choose our baseline assumption: acquiring information upon the need to change prices. Afrouzi et al. (2024) adopt the same timing structure and argue that it is more efficient for firms. Second, under the alternative timing assumption, the channel we emphasize would still operate; however, it would introduce an additional mechanism, as we discuss below. Therefore, our preferred specification can be viewed as a modeling choice that helps isolate the relevant channel we highlight, which operates through general equilibrium forces. Third, the information acquisition behavior implied by our preferred timing assumption is consistent with the evidence presented in Section 5.¹¹ Having said this, we believe that empirically verifying whether firms acquire information before or after learning that they will adjust prices remains an important question for future research.

We now present the results from a model that adopts the alternative timing assumption. Formally, the optimal amount of information is given by:

$$\kappa_{ki}^* = \arg \min_{\kappa_{ki}} (1 - \theta_k) E \left[\frac{\epsilon - 1}{2} (p_{ki}^\diamond - p_{ki}^*)^2 \mid s_{ki}^* \right] + \lambda f(\kappa_{ki}). \quad (3.7)$$

This expression is similar to firms' optimal information acquisition in the baseline model, Equation (2.19). This similarity implies that the channels we highlight continue to operate under the alternative timing assumption.

The difference is that the return to information is discounted by the probability of price adjustment, $1 - \theta_k$. Intuitively, when nominal rigidity is high, the expected return to information acquisition is lower ex ante, reducing firms' incentives to acquire information—a direct mechanism. This mechanism acts as an offsetting force to the general equilibrium forces emphasized in the baseline model and modifies the characterization of optimal policy.

Proposition 6 in Appendix B.1 establishes the condition for optimal monetary policy

¹¹On the other hand, Fabiani et al. (2006) use Euro-area firm-level survey data and report that firms assess their pricing decisions more often than they actually change prices: the median firm considers whether to change its prices about twice a year but ultimately adjusts them only once a year.

under the alternative timing assumption, which results in a different—and more restrictive—threshold.

4 Dynamic Model

In this section, we extend our benchmark model into a dynamic multi-sector framework under rational inattention with an infinite horizon: time $t = 1, 2, \dots, \infty$. Sectoral shocks are dynamic, and households, firms, and the central bank solve optimization problems dynamically.

Timing of the Model The setup and timing of the dynamic model are as follows. At the beginning of the world ($t = 0$), the central bank chooses an optimal inflation stabilization rule that minimizes ex-ante expected social welfare loss. Each subsequent period $t > 0$ consists of four stages, analogous to the stages in the baseline model.

In stage one, firms make a state-contingent information acquisition plan. Specifically, firms decide the amount of information to acquire, $\kappa_{ki,t}$, in a subsequent stage at time t , depending on being able to reset prices or not. In stage two, nature draws unobserved sectoral productivity shocks $\{a_{1,t}, a_{2,t}\}$. In the same stage, Calvo uncertainty—whether firms can reset prices or not—is resolved. In stage three, price-resetting firms acquire information according to their plans made in stage two. The cost of information acquisition is paid in this stage. Firms make price-setting plans (i.e., goods-level supply curves) conditional on their information sets. In stage four, the representative household observes the state of the economy and makes consumption and labor decisions while the goods and labor markets clear. The economy then moves to time $t + 1$, and the same four stages repeat.

An important feature of this dynamic model, absent in the baseline static model, is that firms' optimal information acquisition and pricing decisions are dynamic: they consider the probability of not being able to reset prices in the future. We focus on these two problems and relegate the remaining aspects of the model to Appendix E, as they largely parallel the static setting with the addition of a time t subscript.

4.1 Household's Optimization

We assume households have complete information. The representative household solves a lifetime utility maximization problem:

$$\max_{C_t, B_{t+1}, L_{t,1}, L_{t,2}} E_0 \sum_{t=0}^{\infty} \beta^t \left(\log(C_t) - \sum_{k=1}^2 L_{k,t} \right), \quad (4.1)$$

$$\text{s.t.} \quad P_t C_t + Q_t B_{t+1} = B_t + \sum_{k=1}^2 W_{k,t} L_{k,t} + \sum_{k=1}^2 \Pi_{k,t} + T_t, \quad (4.2)$$

where B_t denotes the amount of government bonds purchased by the household at time t , Q_t is the price of the bond, and β is the utility discount factor.

The household's optimization problem leads to the dynamic version of the labor supply equation and goods demand equations discussed in Section 2.1. Additionally, in this dynamic setting, the household's intertemporal optimization results in a consumption Euler equation. The details are provided in Appendix E.

4.2 Firms' Optimization

As in the static model, we assume a continuum of monopolistically competitive firms in each sector k . Each firm i produces output using the production function $Y_{ki,t} = e^{a_{k,t}} L_{ki,t}$, where $a_{k,t}$ is the log productivity in sector k . We assume $a_{k,t}$ evolves according to an autoregressive process: $a_{k,t} = \rho a_{k,t-1} + \epsilon_{k,t}$, where $\epsilon_{k,t}$ is an i.i.d. shock in sector k .

Firms set prices optimally, subject to information constraints and nominal rigidity.

Price Setting Given its information set, a re-optimizing firm chooses an optimal price to maximize the *net present value* of expected profits. Formally, firms set prices optimally according to:

$$P_{ki,t}^* = \arg \max_{P_{ki,t}} \sum_{h=0}^{\infty} \theta_k^h E_{ki,t} \left\{ Q_{t,t+h} \left(P_{ki,t} Y_{ki,t+h|t} - W_{k,t+h} L_{ki,t+h|t} \right) \mid I_{ki,t} \right\}, \quad (4.3)$$

where $E_{ki,t}$ denotes the firm's expectation conditional on its information at period t , $Q_{t,t+h} \equiv \beta^h \frac{C_t}{C_{t+h}} \frac{P_t}{P_{t+h}}$ is the stochastic discount factor, and $I_{ki,t}$ represents the information set which will be described shortly. Here, $Y_{ki,t+h|t}$ and $L_{ki,t+h|t}$ denote the output and labor in period $t+h$ for a firm ki that last reset its price in period t . The term θ_k^h in the objective func-

tion reflects the probability that the firm will not be able to reset its price in the future. Therefore, $P_{ki,t}^*$ remains a relevant price h periods ahead with probability θ_k^h .

Similar to the static setting, we define the relative productivity $\xi_t \equiv \frac{1}{2}(a_{2,t} - a_{1,t}) = \rho \xi_{t-1} + u_t$, where $u_t \equiv \frac{1}{2}(\epsilon_{2,t} - \epsilon_{1,t})$ follows a normal distribution with a mean zero and variance σ_u^2 . However, unlike the static model where the current period's shock is a state variable, in the dynamic setting, the unknown states include the whole history of shocks. For a length of truncation L , the state space is

$$u^t \equiv (u_t, u_{t-1}, \dots, u_{t-(L-1)})' \in \mathbb{R}^L. \quad (4.4)$$

The law of motion of u^t can be written compactly as $u^t = Au^{t-1} + Qu_t$, where $A \in \mathbb{R}^{L \times L}$ and $Q \in \mathbb{R}^{L \times 1}$ are defined in equation (E.17) in Appendix E.

Afrouzi et al. (2024) proved a theorem showing that, in a dynamic rational inattention setting, firms choose an optimal information structure of the following form¹²:

$$s_{ki,t} = \mathbf{y}_{ki}' u^t + e_{ki,t}, \quad \text{with } e_{ki,t} \sim N(0, \sigma_{e,ki,t}^2), \quad (4.5)$$

where $\mathbf{y}_{ki} \in \mathbb{R}^{L \times 1}$ denotes the loading factor that firm i in sector k optimally chooses for its signal structure. The firm's information set is:

$$I_{ki,t} = s_{ki}^t \equiv (s_{ki,t}, s_{ki,t-1}, \dots, s_{ki,0}), \quad (4.6)$$

where s_{ki}^t represents the sequence of all signals received by firm i up to period t .

Information Acquisition Conditional on resetting prices, a firm ki 's optimal amount of information acquisition solves the following problem:

$$\max_{\Sigma_{ki,t|t} \in \mathcal{S}_+^L} E \left[\sum_{h=0}^{\infty} \theta_k^h Q_{t,t+h} \left(P_{ki,t}^* Y_{ki,t+h|t} - W_{k,t+h} L_{ki,t+h|t} \right) \mid I_{ki,t} \right] - \frac{\lambda}{2} \log_2 \left(\frac{|\Sigma_{ki,t|t-1}|}{|\Sigma_{ki,t|t}|} \right), \quad (4.7)$$

¹²Afrouzi et al. (2024) demonstrates that firms' optimal choice of the number of signals does not exceed the number of actions (excluding the attention choice) for each firm. This implies that firms choose a single signal in our setting.

subject to the optimal price setting rule (4.3), and the constraints on the prior and the posterior uncertainty about u^t :

$$\Sigma_{ki,t+1|t} = A\Sigma_{ki,t|t}A' + QQ', \quad \forall t \geq 0, \quad (4.8)$$

$$\Sigma_{ki,t|t-1} - \Sigma_{ki,t|t} \succeq 0, \quad \forall t \geq 0, \quad (4.9)$$

$$0 \prec \Sigma_{ki,0|-1} \prec \infty \quad \text{given}. \quad (4.10)$$

Here, S_+^L denotes the L -dimensional symmetric positive semidefinite cone, $|\cdot|$ represents the determinant operator, $\succeq$ denotes positive semidefiniteness, $\Sigma_{ki,t|t-1} \equiv \text{var}(u^t | s_{ki}^{t-1})$, $\Sigma_{ki,t|t} \equiv \text{var}(u^t | s_{ki}^t)$, and $\Sigma_{ki,0|-1}$ is the prior variance-covariance matrix in the initial period. Equation (4.8) describes the formation of prior uncertainty and equation (4.9) represents a non-forgetting constraint.

Note that, as discussed previously, the timing assumption of this dynamic model is kept as close as possible to the timing in the baseline static model. Specifically, we assume that firms' decision to acquire information or not at time t is made after the resolution of the *current pricing* uncertainty at time t . In fact, the static model is a special case of the dynamic model. This can be seen from Equation (4.7): when h is forced to take the value 0—the case when only current dividends matter—the objective function collapses to its counterpart in the static model, Equation (2.12).

Interestingly, additional terms appear in this dynamic setting. Similar to firms' dynamic price setting, firms' information acquisition decisions take into account the possibility of not being able to adjust prices in the future. The term θ_k^h in the objective function reflects the probability that the firm will not be able to reset its price in the future. Notably, these *dynamic future pricing* uncertainties are not resolved at the time the information acquisition decisions are made.

These additional terms introduce an additional channel. With more rigid prices, the prices firms set today are likely to last longer. Consequently, the marginal benefit of acquiring information increases with price rigidity due to *dynamic future pricing* uncertainty. This reinforces the general equilibrium channel we emphasize in Proposition 1.

Following Afrouzi et al. (2024), we have assumed a linear attention cost function, $f(\kappa_{ki,t}) = \kappa_{ki,t}$, in this dynamic setting. This assumption simplifies the solution of the model, as it implies that prior uncertainty $\Sigma_{ki,t|t-1}$ does not affect the choice of the optimal $\Sigma_{ki,t|t}$ when firms come to reset prices.

4.3 Central Bank

Following [Ou et al. \(2021\)](#), the second-order approximation of the representative consumer's period welfare loss can be expressed as:

$$E\mathbb{L} = E \left[\sum_{k=1}^2 (\epsilon n_k \text{var}_i(p_{ki,t})) + x_t^2 + \eta^{-1} \sum_{k=1}^2 n_k \tilde{y}_{R,kt}^2 \right], \quad (4.11)$$

where $E\mathbb{L}$ represents the expected welfare loss, x_t is the output gap, and $\tilde{y}_{R,kt}$ is the relative output gap defined as $\tilde{y}_{R,kt} = (y_{k,t} - y_t) - (y_{k,t}^N - y_t^N)$, with $y_{k,t}^N$ and y_t^N being the natural level of sectoral output and aggregate output, respectively, under full information and flexible prices.

Similar to the static model, we decompose the price dispersions into two components: the dispersed-belief price dispersion caused by limited attention and the Calvo price dispersion caused by nominal rigidity. Specifically,

$$E[\text{var}_i(p_{ki,t})] = E \left[\frac{1}{1 - \beta\theta_k} \left[(1 - \theta_k) \int_i (p_{ki,t}^* - p_{k,t}^*)^2 di + \frac{\theta_k}{1 - \theta_k} \pi_{k,t}^2 \right] \right],$$

where $\pi_{k,t} \equiv p_{k,t} - p_{k,t-1}$ denotes inflation in sector k .

In this dynamic model, the OPI stabilization policy is formally defined as:

$$\min_{\omega_1, \omega_2} E\mathbb{L}$$

with $\mathbb{L}$ defined according to (4.11), subject to:

$$\omega_1 \pi_{1,t} + \omega_2 \pi_{2,t} = 0, \quad (4.12)$$

along with other equilibrium conditions. Details are provided in [Appendix E.1](#).

The linearization and characterization of the equilibrium in this model are similar to those in the baseline model. Therefore, these details are delegated to [Appendix E](#). The next section presents the calibration and results.

4.4 Results

We calibrate the dynamic model to address the same research question: In an economy with two sectors that are symmetric except for a heterogeneous degree of nominal rigidi-

ties (i.e., $\theta_1 \neq \theta_2$), does $\theta_1 < \theta_2$ imply $\omega_1^* < \omega_2^*$?

Note that the channel we emphasized in the static model is unambiguously present in the dynamic model. However, the richness of the dynamic setting can alter the quantitative relevance of our channel, increase the quantitative relevance of conventional channels, and potentially introduce additional mechanisms that are omitted in a stylized static setting.

The purpose of addressing the same research question in the dynamic setting is not a quantitative one: it is not intended to quantitatively characterize each potential channel since the dynamic model is also stylized. Instead, the objective is to illustrate whether the channel established in the static setting becomes irrelevant in the dynamic setting or is easily dominated by competing channels.

To demonstrate the relevance of the mechanism we emphasize, it is sufficient to show that our main result carries over to the dynamic setting under standard calibrations.

The parameter calibrations are as follows. The two sectors are identical except for nominal rigidities. We set sector sizes to $n_1 = n_2 = 0.5$, and the elasticity of substitution between sectors is $\eta = 1$, while the elasticity of substitution within a sector is $\epsilon = 11$. The utility discount factor is set to $\beta = 0.99$. The information processing cost $\lambda = 0.036$ is calibrated to match a Kalman gain of 0.3, as estimated in [Coibion and Gorodnichenko \(2012\)](#). Finally, we set the nominal rigidities to $\theta_1 = 0.65$ and $\theta_2 = 0.75$, which correspond to the degrees of nominal rigidity reported in [Gorodnichenko and Weber \(2016\)](#) for the manufacturing and service sectors in the United States, respectively. In terms of the shock process, we consider both the case of i.i.d. shocks and the case of persistent shocks ($\rho = 0.5$), with the same variance ($\sigma_u^2 = 0.02$) across sectors. The results are qualitatively unchanged.

Under these calibrations, with $\theta_1 = 0.65 < \theta_2 = 0.75$, we find $\omega_1^* > \omega_2^*$. More specifically, $\frac{\omega_1^*}{\omega_1^* + \omega_2^*}$ equals 0.95 and 0.6 for the cases with i.i.d. and persistent shocks, respectively. This result is summarized in Column (4) of Table [1](#).

We conclude that the results in this section provide evidence supporting the potential relevance of the proposed mechanism for policy analysis. Future research exploring the optimal monetary policy within a more comprehensive quantitative framework that incorporates this mechanism could prove fruitful.

5 Empirical Evidence

The key to our findings is the endogenous relationship between attention and nominal rigidities, as highlighted in Proposition 1. Specifically, firms in sectors with more flexible prices pay less attention to macroeconomic conditions. In this section, we provide direct empirical evidence supporting this mechanism.

Empirical Proxy for Firms’ Attention to Macroeconomic Conditions Following Song and Stern (2022) and Flynn and Sastry (2022), we use the frequencies of macroeconomic keywords in firms’ reports as a proxy for firms’ attention. Previous literature has demonstrated that these measures serve as sensible proxies for firms’ attention. Song and Stern (2022) and Flynn and Sastry (2022) focused on the cyclicalities of the constructed attention measures, with Song and Stern (2022) documenting heterogeneous effects of monetary policy on firms’ market values depending on their attention to macroeconomic conditions. In this paper, we explore the correlation between firms’ attention and the frequency of price adjustment (FPA).

Our text data are derived from U.S. public companies’ annual Form 10-K filings.¹³ All companies whose securities are traded on a U.S. exchange and are subject to Section 13 or 15(d) of the U.S. Securities Exchange Act of 1934 must file these reports. As summarized by SEC (2011), “Among other things, the 10-K offers a detailed picture of a company’s business, the risks it faces, and the operating and financial results for the fiscal year. Company management also discusses its perspective on the business results and what is driving them.” These reports therefore contain valuable information about what firms pay attention to.

We use the frequency of macroeconomic keywords that appear in firms’ reports as a proxy for the amount of attention (i.e., the amount of information) that firms allocate to the associated macroeconomic topic or condition. To calculate word counts, we employ the method proposed by Loughran and McDonald (2011). This process involves parsing the filings by removing numbers and “stop words,” such as connectors, and mapping each word to a dictionary containing all words in the sample. Subsequently, we apply the same topics and keywords (see Table 3) used by Song and Stern (2022) to construct measures of attention intensity. Specifically, attention intensity $\hat{\kappa}_{i,t}^j$ represents the number

¹³The body of the 10-K filings is retrieved from the Notre Dame Software Repository for Accounting and Finance (SRAF). The SRAF is a repository of financial and accounting data maintained by the University of Notre Dame, accessible at: <https://sraf.nd.edu/data/stage-one-10-x-parse-data/>.

of times topic j is mentioned divided by the total number of words in the filing by a firm i at time t :

$$\hat{\kappa}_{i,t}^j = \frac{\text{Total topic } j \text{ words}_{i,t}}{\text{Total words}_{i,t}}.$$

Our empirical assumption is that, *conditional on* a firm using the selected macroeconomic keywords in its report, the frequency of these keywords is proportional to the firm's attention to the corresponding macroeconomic condition. This proxy aligns with the intensity measure of attention defined by Song and Stern (2022).

At this point, a clarification of the identification assumption is useful. Specifically, we acknowledge the possibility that companies might be highly attentive to macroeconomic conditions but choose a set of words different from our selected keywords due to their writing style. Consequently, ranking companies that do not use the selected keywords as paying less attention than a company that mentions a keyword once would be inappropriate. Therefore, we exclude these observations and rely on the frequency of macroeconomic keywords as a proxy for attention only among firms that use these keywords in their reports. Additionally, our theory relies on the negative relationship between the attention of price-resetting firms and the nominal rigidities in their sector. Excluding firms that do not mention macroeconomic conditions helps restrict the sample to the relevant subset of firms.

Merging the Attention Data with Frequency of Price Adjustment Data We use the FPA dataset constructed by Pasten et al. (2020). Utilizing micro-level data underlying the construction of the Producer Price Index, the authors calculate the frequency of price adjustments at the NAICS 6-digit level. This sectoral-level data is suitable for our analysis, as our model relies on the correlation between FPA and attention at the *sectoral* level.

To merge the two datasets, we aggregate the firm-level attention data to the NAICS 6-digit sector level by calculating the sales-weighted average of firms within each sector.¹⁴

In the final dataset, for each topic j , we observe $\hat{\kappa}_{k,j}$, the average attention paid by firms in sector k , and FPA_k , the frequency of price adjustment in the same sector. A more detailed description of the data construction process can be found in Appendix F.

Empirical Results To investigate the relationship between firms' attention to macroeconomic conditions and the degree of nominal rigidity, we estimate the following regres-

¹⁴We first compute the average attention and average sales (from Compustat) for each firm. The sectoral attention data is then calculated as the sales-weighted average of firms within each sector.

sion:

$$100 \times \log(\hat{\kappa}_{k,j}) = \alpha + \beta \times FPA_k + FE + u_{k,j}, \quad (5.1)$$

where FE represents fixed effects.¹⁵

Table 2 reports the estimation results. The parameter of interest is β , which is associated with the variable FPA_k . Across all specifications we consider, the estimated $\hat{\beta}$ is negative, consistent with the model. This implies that in sectors with more flexible prices (higher FPA_k), firms pay less attention to macroeconomic conditions, as measured by the frequency of macroeconomic keywords in firms' reports. Column (1) reports the result of estimating the baseline model (5.1) without including fixed effects. In Column (2), we control for topic fixed effects. In Column (3), we include topic-by-sector (NAICS 2-digit) fixed effects to address potential confounding factors that are common at the NAICS 2-digit level. Across all specifications, the main finding remains robust.

To address the potential concern that the findings might be driven by extreme observations—namely, firms that mentioned macroeconomic keywords only once—we re-estimate the empirical models after excluding these firms. Table 4 demonstrates that the empirical results are robust to the exclusion of such extreme observations.

Additionally, the baseline pooled regression assumes a homogeneous relationship across topics. To further explore potential heterogeneity, Table 5 reports regression results by topic. A negative relationship is observed across all topics except for oil. For the topics related to consumption and inflation—arguably the most relevant variables in our model—the negative relationship is statistically significant. Interestingly, for the topic related to oil, a positive relationship is observed.¹⁶

¹⁵Within each sector, there are nine measures of attention corresponding to different topics, along with one measure of sector-specific FPA. Equation 5.1 is estimated using pooled OLS.

¹⁶Explaining the heterogeneous relationship across topics is left for future research, as it lies beyond the scope of this paper.

Table 2: Regression Results

	(1)	(2)	(3)
$\hat{\beta}$	-33.19*** (8.18)	-26.97*** (6.44)	-27.58*** (6.92)
Topic FE	NO	YES	NO
Sector $\times$ Topic FE	NO	NO	YES
Observations	3303	3303	3303
Adjusted R-squared	0.005	0.384	0.421

Note: This table report the estimated $\hat{\beta}$ from $100 \times \log(\hat{k}_{k,j}) = \alpha + \beta \times FPA_k + FE + u_{k,j}$, where FPA indicates the frequency of price adjustment. $\hat{k}_{k,j}$ measures the amount of information. Sector FE refers to the 2-digit NAICS sector, and Topic FE refers to the group of keywords used to measure attention. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6 Conclusion

Most central banks in advanced economies stabilize the Consumer Price Index (CPI), which is defined as the price index weighted by consumption shares (or sector sizes). The existing academic literature emphasizes the importance of accounting for cross-sector heterogeneities. The prevailing view suggests that the optimal price index should assign greater weight to prices in sectors with relatively stickier prices. This paper challenges this policy prescription by incorporating rational inattention into existing frameworks.

We demonstrate that with endogenous information acquisition, firms in sectors with more flexible prices gather less information. In other words, sectors with lower degrees of nominal frictions are endogenously associated with more pronounced information frictions. These heightened information frictions increase price dispersion among price-resetting firms in flexible price sectors, thereby exacerbating social welfare losses. As a result, our model introduces a channel through which optimal monetary policy tends to prioritize stabilizing prices in sectors with more flexible prices. We analytically show that when information acquisition is sufficiently costly, this channel dominates, leading to an optimal monetary policy that places greater emphasis on stabilizing prices in sectors with more flexible prices.

We also compare the performance of CPI stabilization with that of the Stabilized Price Index (SPI) policy. The SPI represents the optimal price index derived from an equivalent

model under the assumption of full information. Our findings indicate that stabilizing the CPI yields better outcomes in terms of social welfare.

To support these findings, we provide empirical evidence that substantiates the core mechanism, using the frequency of macroeconomic keywords in firms' reports as a proxy for firms' attention to unobserved macroeconomic conditions. Consistent with the mechanism in our model, we show that firms in sectors with more frequent price adjustments exhibit lower attention to macroeconomic conditions.

Our findings convey a broader message: an optimal policy derived from a stylized model *might* be misleading if omitted frictions correlate with the existing frictions. This critique also applies to our model. Specifically, in our framework, the optimal monetary policy incentivizes firms in sectors with flexible prices to acquire limited information. Within the model, such an allocation is optimal because the central bank stabilizes the actions of these partially informed firms. However, in practice, firms' information choices may interact with other sources of distortion, potentially introducing additional costs related to information acquisition. Investigating this possibility could offer a promising direction for future research.

In terms of policy relevance, our theoretical discussions and quantitative explorations suggest that the channel we emphasize, which operates in opposition to conventional channels, is qualitatively present and potentially quantitatively relevant. However, the models presented, including the dynamic model, are stylized and may not be well-suited for assessing the *quantitative* significance of this mechanism in comparison to other mechanisms. Future research exploring optimal monetary policy within a more comprehensive quantitative framework that incorporates this mechanism could prove fruitful.

References

- Afrouzi, H., J. P. Flynn, and C. Yang (2024). What can measured beliefs tell us about monetary non-neutrality? Technical report, National Bureau of Economic Research. [1](#), [3.6](#), [4.2](#), [12](#), [4.2](#), [E](#)
- Afrouzi, H. and C. Yang (2021). Dynamic rational inattention and the phillips curve. [1](#), [E](#), [E.2](#), [E.2](#), [E.2](#)
- Alvarez, F., F. Lippi, and A. Oskolkov (2022). The macroeconomics of sticky prices with generalized hazard functions. *The Quarterly Journal of Economics* 137(2), 989–1038. [1](#)
- Anand, R., E. S. Prasad, and B. Zhang (2015). What Measure of Inflation should a Developing Country Central Bank Target? *Journal of Monetary Economics* 74(C), 102–116. [1](#)
- Andrade, P., R. K. Crump, S. Eusepi, and E. Moench (2016). Fundamental Disagreement. *Journal of Monetary Economics* 83, 106–128. [1](#)
- Angeletos, G.-M., L. Iovino, and J. La’o (2016). Real Rigidity, Nominal Rigidity, and the Social Value of Information. *American Economic Review* 106(1), 200–227. [1](#)
- Aoki, K. (2001, August). Optimal Monetary Policy Responses to Relative-price Changes. *Journal of Monetary Economics* 48(1), 55–80. [1](#)
- Basu, S. and P. D. Leo (2016, March). Should Central Banks Target Investment Prices? Boston College Working Papers in Economics 910. [1](#)
- Benigno, P. (2004, July). Optimal Monetary Policy in a Currency Area. *Journal of International Economics* 63(2), 293–320. [1](#)
- Broer, T., A. Kohlhas, K. Mitman, and K. Schlafmann (2022). Expectation and wealth heterogeneity in the macroeconomy. [3](#)
- Camous, A. and D. Matveev (2023). Central bank strategic communication and the dynamics of reputation. [1](#)
- Carvalho, C. (2006). Heterogeneity in price stickiness and the real effects of monetary shocks. *Frontiers in Macroeconomics* 2(1). [1](#)

- Carvalho, C., N. A. Dam, and J. W. Lee (2020). The cross-sectional distribution of price stickiness implied by aggregate data. *Review of Economics and Statistics* 102(1), 162–179. [1](#)
- Carvalho, C. and F. Schwartzman (2015). Selection and monetary non-neutrality in time-dependent pricing models. *Journal of Monetary Economics* 76, 141–156. [1](#)
- Coibion, O. and Y. Gorodnichenko (2011). Strategic interaction among heterogeneous price-setters in an estimated dsge model. *Review of Economics and Statistics* 93(3), 920–940. [1](#)
- Coibion, O. and Y. Gorodnichenko (2012). What Can Survey Forecasts Tell Us about Information Rigidities? *Journal of Political Economy* 120(1), 116–159. [1](#), [3.3](#), [2](#), [3](#), [4](#), [4.4](#)
- Coibion, O. and Y. Gorodnichenko (2015). Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. *American Economic Review* 105(8), 2644–78. [1](#)
- Edmond, C., V. Midrigan, and D. Y. Xu (2018). How costly are markups? [3.3](#)
- Ellison, M. and A. Macaulay (2021). A rational inattention unemployment trap. *Journal of Economic Dynamics and Control* 131, 104226. [3](#)
- Fabiani, S., M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, F. Martins, T. Matha, R. Sabbatini, H. Stahl, et al. (2006). What firms’ surveys tell us about price-setting behavior in the euro area. *International Journal of Central Banking*. [11](#)
- Fang, Y., Z. Huo, M. Li, and J. Wu (2024). Inattentive network: Evidence and theory. [1](#), [3.2](#)
- Flynn, J. P. and K. Sastry (2022). Attention cycles. *Available at SSRN* 3592107. [1](#), [5](#)
- Gaballo, G. (2016). Rational inattention to news: The perils of forward guidance. *American Economic Journal: Macroeconomics* 8(1), 42–97. [3](#)
- Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press. [3.3](#)
- Gautier, E. and H. L. Bihan (2018). Shocks vs menu costs: Patterns of price rigidity in an estimated multi-sector menu-cost model. *The Review of Economics and Statistics*, 1–45. [1](#)

- Gorodnichenko, Y. and M. Weber (2016, January). Are Sticky Prices Costly? Evidence from the Stock Market. *American Economic Review* 106(1), 165–199. 1, 4.4
- Hobijn, B. and F. Nechio (2019). Sticker Shocks: Using VAT Changes to Estimate Upper-Level Elasticities of Substitution. *Journal of the European Economic Association* 17(3), 799–833. 3.3
- Höynck, C., M. Li, and D. Zhang (2022). The markup elasticity of monetary non-neutrality. *Available at SSRN* 4283059. 1
- Huang, K. X. and Z. Liu (2005, November). Inflation Targeting: What Inflation Rate to Target? *Journal of Monetary Economics* 52(8), 1435–1462. 1
- Ilut, C. and R. Valchev (2023). Economic agents as imperfect problem solvers. *The Quarterly Journal of Economics* 138(1), 313–362. 3
- Lan, T., L. Li, and M. Li (2024). Cyclical inequality in the cost of living and implications for monetary policy. 1
- La’O, J. and A. Tahbaz-Salehi (2022). Optimal monetary policy in production networks. *Econometrica* 90(3), 1295–1336. 1, 3.2, 3.3
- Li, M. and H.-M. Wu (2016). Optimal Monetary Policy with Asymmetric Shocks and Rational Inattention. *mimeo*. 1, D.1
- Loughran, T. and B. McDonald (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *Journal of Finance* 66(1), 35–65. 1, F.2
- Loughran, T. and B. McDonald (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research* 54(4), 1187–1230. F.2
- Luo, Y. (2008). Consumption dynamics under information processing constraints. *Review of Economic dynamics* 11(2), 366–385. 3
- 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. 3
- Luo, Y., J. Nie, and P. Yin (2022). Attention allocation and heterogeneous consumption responses. *Working Paper*. 3

- Maćkowiak, B., F. Matějka, and M. Wiederholt (2023). Rational inattention: A review. *Journal of Economic Literature* 61(1), 226–273. [3](#)
- Mackowiak, B., E. Moench, and M. Wiederholt (2009). Sectoral Price Data and Models of Price Setting. *Journal of Monetary Economics* 56(S), 78–99. [2.3](#)
- Mackowiak, B. and M. Wiederholt (2009). Optimal Sticky Prices under Rational Inattention. *American Economic Review* 99(3), 769–803. [1](#), [2.3](#), [B.2](#), [C.2](#), [E](#)
- Maćkowiak, B. and M. Wiederholt (2015). Business cycle dynamics under rational inattention. *The Review of Economic Studies* 82(4), 1502–1532. [3](#)
- Mankiw, N. G. and R. Reis (2002). Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics* 117(4), 1295–1328. [1](#)
- Mankiw, N. G. and R. Reis (2003, September). What Measure of Inflation Should a Central Bank Target? *Journal of the European Economic Association* 1(5), 1058–1086. [1](#)
- Marcet, A. and K. J. Singleton (1999). Equilibrium asset prices and savings of heterogeneous agents in the presence of incomplete markets and portfolio constraints. *Macroeconomic Dynamics* 3(2), 243–277. [22](#)
- Matějka, F. (2016). Rationally inattentive seller: Sales and discrete pricing. *The Review of Economic Studies* 83(3), 1125–1155. [3](#)
- Matějka, F. and A. McKay (2012). Simple market equilibria with rationally inattentive consumers. *American Economic Review* 102(3), 24–29. [3](#)
- Melosi, L. (2016). Signalling Effects of Monetary Policy. *The Review of Economic Studies* 84(2), 853–884. [1](#)
- Miao, J., J. Wu, and E. R. Young (2022). Multivariate rational inattention. *Econometrica* 90(2), 907–945. [3](#)
- Miao, J. and B. Zhang (2024). Linear quadratic approximation of rationally inattentive control problems. *Macroeconomic Dynamics* 28(6), 1371–1393. [3](#)
- Mondria, J. (2010). Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory* 145(5), 1837–1864. [3](#), [D.1](#)

- Nakamura, E. and J. Steinsson (2008). Five facts about prices: A reevaluation of menu cost models. *The Quarterly Journal of Economics* 123(4), 1415–1464. [1](#)
- Nimark, K. (2008). Dynamic Pricing and Imperfect Common Knowledge. *Journal of monetary Economics* 55(2), 365–382. [1](#)
- Olivi, A., V. Sterk, and D. Xhani (2024). Optimal monetary policy during a cost of living crisis. Technical report, mimeo, UCL. [1](#)
- Ou, S., D. Zhang, and R. Zhang (2021). Information frictions, monetary policy, and the paradox of price flexibility. *Journal of Monetary Economics* 120, 70–82. [1](#), [2.3](#), [2.4](#), [4.3](#), [18](#)
- Ou, S., D. Zhang, and R. Zhang (2022). The return of greenspan: Mumbling with great incoherence. *Available at SSRN* 4085073. [1](#)
- Paciello, L. and M. Wiederholt (2014). Exogenous Information, Endogenous Information, and Optimal Monetary Policy. *Review of Economic Studies* 81(1), 356–388. [1](#), [3.1](#), [C.3](#)
- Pasten, E. and R. Schoenle (2016). Rational inattention, multi-product firms and the neutrality of money. *Journal of Monetary Economics* 80, 1–16. [3](#)
- 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. [1](#), [1](#)
- Rubbo, E. (2022). Networks, phillips curves, and monetary policy. *manuscript, Harvard University*. [1](#)
- SEC (2011, January). Investor Bulletin: How to Read a 10-K. Technical report, Securities and Exchange Commission Office of Investor Education and Advocacy. [5](#)
- Sims, C. A. (2003). Implications of Rational Inattention. *Journal of monetary Economics* 50(3), 665–690. [1](#), [3](#)
- Song, W. and S. Stern (2022, January). Firm Inattention and the Efficacy of Monetary Policy: A Text-Based Approach. Staff Working Papers 22-3, Bank of Canada. [1](#), [5](#), [1](#), [F.2](#)
- Stevens, L. (2020). Coarse pricing policies. *The Review of Economic Studies* 87(1), 420–453. [3](#)
- Turen, J. (2023). State-dependent attention and pricing decisions. *American Economic Journal: Macroeconomics* 15(2), 161–189. [3](#)

- Tutino, A. (2013). Rationally inattentive consumption choices. *Review of Economic Dynamics* 16(3), 421–439. 3
- Van Nieuwerburgh, S. and L. Veldkamp (2010). Information acquisition and underdiversification. *The Review of Economic Studies* 77(2), 779–805. 3
- Veldkamp, L. L. (2011). *Information choice in macroeconomics and finance*. Princeton University Press. 3
- Woodford, M. (2003). *Interest and prices: Foundations of a theory of monetary policy*. princeton university press. 2.4
- Yang, C. (2022). Rational inattention, menu costs, and multi-product firms: Micro evidence and aggregate implications. *Journal of Monetary Economics* 128, 105–123. 1
- Yin, P. (2021). Optimal attention and heterogeneous precautionary saving behavior. *Journal of Economic Dynamics and Control* 131, 104230. 3
- Zhang, D. (2022). Term structure, forecast revision, and the signaling channel of monetary policy. *Journal of the European Economic Association* 20(4), 1522–1553. 1
- Zorn, P. (2020). Investment under rational inattention: Evidence from us sectoral data. 3

A Figures and Tables

Table 3: Macroeconomics Topics and Keywords

Topic	Keywords
General	economic conditions
Output	GDP , economic growth, macroeconomic condition, construction spending, national activity, recession
Employment	national activity, JOLTS , labor market, jobless claims, jobs report, non farm payroll, adp employment report, employment cost index
Consumption	consumer confidence, consumer credit, consumer sentiment, durable goods, personal income, retail sales
Investment	business inventories, manufacturing survey, factory orders, business outlook survey, manufacturing index, industrial production, business optimism, wholesale trade
FOMC	FOMC , monetary policy, quantitative easing
Housing	home sales, home prices, housing starts, housing market
Inflation	price index, price level, consumer price index, CPI , PMI , PPI , disinflationary, hyperinflation, inflation, inflationary, hyperinflationary, disinflation
Oil	oil prices, oil supply, oil demand

Table 4: Regression Results: Robustness Check

	(1)	(2)	(3)
$\hat{\beta}$	-40.05*** (7.31)	-32.72*** (6.38)	-30.83*** (6.80)
Topic FE	NO	YES	NO
Sector * Topic FE	NO	NO	YES
Observations	2583	2583	2583
Adjusted R-squared	0.011	0.251	0.297

Note: This table report the estimated $\hat{\beta}$ from $100 \times \log(\hat{k}_{k,j}) = \alpha + \beta \times FPA_k + FE + u_{k,j}$, where FPA indicates the frequency of price adjustment. $\hat{k}_{k,j}$ measures the amount of information. Sector FE refers to the 2-digit NAICS sector, and Topic FE refers to the group of keywords used to measure attention. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Regression Results by Topic

Topic	No FE	Sector FE	Observations
General	-10.70 (12.28)	-8.56 (13.53)	561
Output	-11.84 (14.71)	-3.80 (15.85)	488
Employment	-44.87** (21.82)	-36.31 (25.37)	254
Consumption	-56.77*** (20.80)	-64.42*** (21.17)	472
Investment	-62.89*** (22.66)	-52.86** (23.76)	219
FOMC	-29.20 (24.28)	-43.23 (29.35)	143
Housing	-38.64 (29.19)	-38.53 (30.83)	296
Inflation	-41.29*** (14.84)	-42.65*** (16.49)	569
Oil	45.48** (22.95)	22.46 (23.42)	301

Note: This table report the estimated $\hat{\beta}_j$ from $100 \times \log(\hat{k}_{k,j}) = \alpha + \beta_j \times FPA_k + \text{Sector FE} + u_{k,j}$, for each top j , where FPA indicates the frequency of price adjustment. $\hat{k}_{k,j}$ measures the amount of information. Sector FE refers to the 2-digit NAICS sector. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B Model: Additional Results and Equations

B.1 Proposition 6

Proposition 6. *Consider the economy with endogenous information acquisition and $\theta_1 < \theta_2$. There exists a threshold value $\bar{K}^{alt}$ such that*

$$\omega_1^* > \omega_2^* \quad \text{if and only if} \quad K^s < \bar{K}^{alt}.$$

That is, the optimal monetary policy attaches a smaller weight to the sector with stickier prices if and only if information acquisition is sufficiently costly.

Proof. See Appendix [D.10](#). □

B.2 Linear-Quadratic Approximation of the Economy

To solve the model analytically, this section log-linearizes agents' optimality conditions. To this end, we make use of the second-order approximation of firms' profit function to solve for the firms' information acquisition problem, and the second-order approximation of the welfare function to solve for the central bank's optimal monetary policy.

Throughout the paper, a variable in lowercase denotes the log deviation of the variable from its initial value in the non-stochastic solution.

Household's Labor Supply The log-linearized version of the household's labor supply [\(2.1\)](#) is:

$$w_k = p + c, \tag{B.1}$$

where

$$p = n_1 p_1 + n_2 p_2. \tag{B.2}$$

Firms' Optimal Price Setting The log-linearized optimal price resetting rule [\(2.7\)](#) is:

$$p_{ki}^* = E[p_{ki}^\diamond \mid I_{ki}], \tag{B.3}$$

where $p_{ki}^\diamond$ denotes the profit-maximizing price of firm ki had this firm fully observed the states of the world ([Mackowiak and Wiederholt 2009](#)):

$$p_{ki}^\diamond = p + x + u_k, \tag{B.4}$$

where $x = y - y^N$ is the output gap, and $y^N \equiv n_1 a_1 + n_2 a_2$ is the natural output, defined as the level of output that would prevail in the absence of nominal and information frictions.¹⁷ n_1 and n_2 indicate the sizes of the two sectors, $u_1 = n_1(a_2 - a_1)$ and $u_2 = n_2(a_1 - a_2)$.

¹⁷Detailed derivation can be found in Appendix [C.1](#).

As explained above, we consider the case where the sizes of the two sectors are the same, i.e., $n_1 = n_2 = 0.5$, to exclude heterogeneity arising from sector sizes and focus our analysis on heterogeneous nominal frictions. As a result:

$$u_1 = -u_2 = u, \quad (\text{B.5})$$

with $u = 0.5(a_2 - a_1)$ capturing the asymmetric nature of sectoral shocks. Under a price stabilization policy, u is sufficient to solve the model.¹⁸ From this point onward, we focus on the relevant unknown state, u , which is assumed to be drawn from a normal distribution:

$$u \sim N(0, \sigma_u^2). \quad (\text{B.6})$$

Given the Calvo nominal rigidities, the linearized aggregate sector prices are:

$$p_k = (1 - \theta_k)p_k^* = (1 - \theta_k) \int_i p_{ki}^* di. \quad (\text{B.7})$$

Therefore, the linearized aggregate supply curves are:

$$p_k = (1 - \theta_k) \int_i E[p + x + u_k \mid I_{ki}] di. \quad (\text{B.8})$$

Optimal Monetary Policy The OPI stabilization policy solves the following linear-quadratic minimization:

$$\min_{\omega} E\mathbb{L}$$

subject to:

- (i) Optimal price setting characterized by (B.3) and (B.4): $p_{ki}^* = E[p + x + u_k \mid I_{ki}]$;
- (ii) Firms' information set (2.17): $I_{ki} = s_{ki}^* = u + e_{ki}$, $e_{ki} \sim N(0, \sigma_{e,k}^2)$;
- (iii) Information acquisition choice (2.19): $\kappa_{ki}^* = \arg \min_{\kappa_{ki}} E \left[\frac{\epsilon-1}{2} (p_{ki}^\diamond - p_{ki}^*)^2 \mid s_{ki}^* \right] + \lambda f(\kappa_{ki})$;

¹⁸This is a known result (see, e.g., [Ou et al. 2021](#)) and can be verified from the solution of the model. All endogenous variables relevant for welfare analysis, including the output gap x , are functions of $u = 0.5(a_2 - a_1)$. However, this result does not imply that *all* endogenous variables can be solved as functions of u . For instance, determining y and y^N would require tracking a_1 and a_2 separately.

- (iv) Aggregation of prices: $p_k = (1 - \theta_k)p_k^*$, $p_k^* = \int p_{ki}^* di$, and $p = 0.5p_1 + 0.5p_2$;
- (v) Definition of the relative price gap: $\tilde{p}_1 - \tilde{p}_2 \equiv p_1 - p_2 - 2u$;
- (vi) The shock process: $u \sim N(0, \sigma_u^2)$;
- (vii) Monetary policy rule (2.23): $\omega p_1 + (1 - \omega)p_2 = 0$.

Note that conditions (i) to (iv) jointly characterize the aggregate supply side of the economy.

B.3 Equilibrium of the Model

For a given monetary policy rule $\{\omega\}$, the equilibrium of the model consists of the household's consumption and labor decisions $\{C_{ki}, \{L_k\}\}$, firms' choice of the signal structure M_{ki}^* and $\Sigma_{e,ki}^*$, the amount of information contained in that signal κ_{ki} , prices P_{ki} , aggregate prices P_k, P , wages $\{W_k\}$, aggregate consumption $\{C_k\}, C$, production Y_{ki} , and aggregate output $\{Y_k\}, Y$, such that:

- (i) Given $\{P_k\}, P, \{W_k\}, Y, \{Y_k\}, \{a_k\}$, each firm $i \in [0, 1]$ in sector k chooses a signal according to (2.11) and a state-contingent information acquisition plan according to (2.15) if allowed to reset prices. After the resolution of Calvo uncertainty, price-resetting firms set the prices for their goods according to (2.7).
- (ii) Given $P_{ki}, \{P_k\}, P, \{W_k\}$, the household supplies labor according to (2.1) and demands goods according to (2.2).
- (iii) Aggregate prices are given by (2.3) and (2.8). Goods markets clear at the individual goods level, $C_{ki} = Y_{ki}$, and at the aggregate level, $C = Y$, $C_k = Y_k \forall k$. Wages $\{W_k\}$ are such that the labor market clears: the aggregate labor supply (2.1) equals the aggregate demand for labor implied by the production function (2.4).

C Appendix for Model Setup and Solution

C.1 Optimal Price Setting

Given the information set of a firm i in sector k , a re-optimizing firm chooses an optimal price to maximize its expected profit:

$$E(\pi_{ki}|I_{ki}) = E(P_{ki}Y_{ki}/P - (1 - \tau)W_k L_{ki}/P | I_{ki})$$

Moreover, from the household's utility maximization problem, we derive:

$$\frac{1}{C} \frac{W_k}{P} - 1 = 0 \Rightarrow W_k = PC \quad (C.1)$$

Thus, the profit function is given by:

$$\pi_{ki} = P_{ki}Y_{ki}/P - (1 - \tau)Ye^{-a_k}Y_{ki} \quad (C.2)$$

Given the optimal consumption solution $C_{k,i} = \left(\frac{P_{k,i}}{P_k}\right)^{-\epsilon} \frac{1}{n_k} C_k$ and $C_k = \left(\frac{P_k}{P}\right)^{-\eta} n_k C$. The firm's profit can be expressed as:

$$\pi_{ki} = \left(\frac{P_{ki}}{P}\right)^{1-\epsilon} \left(\frac{P_k}{P}\right)^{\epsilon-\eta} Y - (1 - \tau)e^{-a_k} \left(\frac{P_{ki}}{P}\right)^{-\epsilon} \left(\frac{P_k}{P}\right)^{\epsilon-\eta} Y^2 \quad (C.3)$$

We rewrite $Y = \bar{Y}e^y$, $P = \bar{P}e^p$, $P_k = \bar{P}e^{p_k}$, $P_{ki} = \bar{P}e^{p_{ki}}$, where a small letter denotes the log-deviation of the variable from its value at the non-stochastic solution (when $a_k = 0$). Now, we denote the profit function v_{ki} which depends on p_{ki} , p_k , p , a_k and y :

$$v_{ki} = e^{(1-\epsilon)p_{ki}} e^{(\epsilon-\eta)p_k} e^{(\eta-1)p} \bar{Y} e^y - (1 - \tau)e^{-a_k} e^{-\epsilon p_{ki}} e^{(\epsilon-\eta)p_k} e^{\eta p} \bar{Y}^2 e^{2y} \quad (C.4)$$

To approximate the equation, we need the following first and second-order derivatives:

$$\frac{\partial v_{ki}}{\partial p_{ki}} = (1 - \epsilon) e^{(1-\epsilon)p_{ki}} e^{(\epsilon-\eta)p_k} e^{(\eta-1)p} \bar{Y} e^y + \epsilon e^{-a_k} e^{-\epsilon p_{ki}} e^{(\epsilon-\eta)p_k} e^{\eta p} \bar{Y}^2 e^{2y} (1 - \tau) \quad (C.5)$$

$$\frac{\partial^2 v_{ki}}{\partial p_{ki}^2} = (1 - \epsilon)^2 e^{(1-\epsilon)p_{ki}} e^{(\epsilon-\eta)p_k} e^{(\eta-1)p} \bar{Y} e^y - \epsilon^2 e^{-a_k} e^{-\epsilon p_{ki}} e^{(\epsilon-\eta)p_k} e^{\eta p} \bar{Y}^2 e^{2y} (1 - \tau) \quad (C.6)$$

$$\frac{\partial^2 v_{ki}}{\partial p_{ki} \partial p_k} = (1 - \epsilon) (\epsilon - \eta) e^{(1-\epsilon)p_{ki}} e^{(\epsilon-\eta)p_k} e^{(\eta-1)p} \bar{Y} e^y + \epsilon (\epsilon - \eta) e^{-a_k} e^{-\epsilon p_{ki}} e^{(\epsilon-\eta)p_k} e^{\eta p} \bar{Y}^2 e^{2y} (1 - \tau) \quad (C.7)$$

$$\frac{\partial^2 v_{ki}}{\partial p_{ki} \partial p} = \eta (1 - \epsilon) e^{(1-\epsilon)p_{ki}} e^{(\epsilon-\eta)p_k} e^{(\eta-1)p} \bar{Y} e^y + (1 + \eta) \epsilon e^{-a_k} e^{-\epsilon p_{ki}} e^{(\epsilon-\eta)p_k} e^{\eta p} \bar{Y}^2 e^{2y} (1 - \tau) \quad (C.8)$$

$$\frac{\partial^2 v_{ki}}{\partial p_{ki} \partial a_k} = -\epsilon e^{-a_k} e^{-\epsilon p_{ki}} e^{(\epsilon-\eta)p_k} e^{\eta p} \bar{Y}^2 e^{2y} (1 - \tau) \quad (C.9)$$

$$\frac{\partial^2 v_{ki}}{\partial p_{ki} \partial y} = (1 - \epsilon) e^{(1-\epsilon)p_{ki}} e^{(\epsilon-\eta)p_k} e^{(\eta-1)p} \bar{Y} e^y + 2\epsilon e^{-a_k} e^{-\epsilon p_{ki}} e^{(\epsilon-\eta)p_k} e^{\eta p} \bar{Y}^2 e^{2y} (1 - \tau) \quad (C.10)$$

Next we evaluate the derivatives above at the non-stochastic solution, i.e. $p_{ki} = 0, p_k = 0, a_k = 0, y = 0$. The employment subsidy τ is chosen to eliminate the inefficiency resulting from the presence of market power of monopolistic competitive firms, i.e. $1 - \tau = \frac{\epsilon-1}{\epsilon}$, implying that $\bar{Y} = 1$. The evaluated derivatives are as follows:

$$\frac{\partial v_{ki}}{\partial p_{ki}}(p_{ki} = 0, p_k = 0, p = 0, y = 0, a_k = 0) = 0, \quad (C.11)$$

$$\frac{\partial^2 v_{ki}}{\partial p_{ki}^2}(p_{ki} = 0, p_k = 0, p = 0, y = 0, a_k = 0) = -(\epsilon - 1), \quad (C.12)$$

$$\frac{\partial^2 v_{ki}}{\partial p_{ki} \partial p_k}(p_{ki} = 0, p_k = 0, p = 0, y = 0, a_k = 0) = 0, \quad (C.13)$$

$$\frac{\partial^2 v_{ki}}{\partial p_{ki} \partial p}(p_{ki} = 0, p_k = 0, p = 0, y = 0, a_k = 0) = \epsilon - 1 \quad (C.14)$$

$$\frac{\partial^2 v_{ki}}{\partial p_{ki} \partial a_k}(p_{ki} = 0, p_k = 0, p = 0, y = 0, a_k = 0) = -(\epsilon - 1) \quad (C.15)$$

$$\frac{\partial^2 v_{ki}}{\partial p_{ki} \partial y}(p_{ki} = 0, p_k = 0, p = 0, y = 0, a_k = 0) = \epsilon - 1 \quad (C.16)$$

Using these results above and together with Equation (C.4), we take second-order approximation to the profit function and obtain

$$v_{ki} \approx -\frac{\epsilon-1}{2} p_{ki}^2 + (\epsilon-1) p_{ki} p - (\epsilon-1) p_{ki} a_k + (\epsilon-1) p_{ki} y + t.i.p.s \quad (C.17)$$

where *t.i.p.s* are terms independent of p_{ki} . Finally, profit-maximizing price $p_{ki}^\diamond$ set by firm i in sector k as implied by the first-order condition $\frac{\partial v_{ki}}{\partial p_{ki}} = 0$ is

$$p_{ki}^\diamond = p + y - a_k \quad (C.18)$$

C.2 Objective Function for Amount of Information Acquisition: Profit Loss

From the inner optimization problem in Equation (2.6), the price set by firm i in sector k is $p_{ki}^* = E[p_{ki}^\diamond | I_{ki}]$, where $p_{ki}^\diamond$ denotes the profit-maximizing price (details can be found above in B.1). Then, we follow Mackowiak and Wiederholt (2009) and define the loss in profits caused by a suboptimal price set by the inattentive decision-maker as

$$\begin{aligned} & v_{ki}(p_{ki}^\diamond, p, y, a_k) - v_{ki}(p_{ki}^*, p, y, a_k) \quad (C.19) \\ &= \frac{1-\epsilon}{2} (p_{ki}^\diamond)^2 + (\epsilon-1) p_{ki}^\diamond p - (\epsilon-1) p_{ki}^\diamond a_k + (\epsilon-1) p_{ki}^\diamond y \\ &\quad - \frac{1-\epsilon}{2} (p_{ki}^*)^2 - (\epsilon-1) p_{ki}^* p + (\epsilon-1) p_{ki}^* a_k - (\epsilon-1) p_{ki}^* y \\ &= \frac{1-\epsilon}{2} [(p_{ki}^\diamond)^2 - (p_{ki}^*)^2] + (\epsilon-1)(p + y - a_k)(p_{ki}^\diamond - p_{ki}^*), \\ &= \frac{\epsilon-1}{2} (p_{ki}^\diamond - p_{ki}^*)^2 \quad (C.20) \end{aligned}$$

where the first equality follows from Equation (C.17) and the second equality follows from Equation (C.18).

C.3 Solution of the Model

The solution for a given choice of price index stabilization rule ω_1 and ω_2 is obtained as follows. First, we normalize the policy rule as follows: $\frac{\omega_1}{\omega_1+\omega_2} p_1 + (1 - \frac{\omega_1}{\omega_1+\omega_2}) p_2 = 0$ and define $\omega = \frac{\omega_1}{\omega_1+\omega_2}$.¹⁹ Second, given the precision of signals, we solve the model to

¹⁹We exclude the parameter region where $\omega_1 + \omega_2 = 0$ throughout the analysis.

obtain the pricing functions governed by the Kalman gains. Then, substitute the pricing functions into the firms' attention choice problem and solve for the optimal attention choice.

Specifically, given the precision of firms' signals, we conjecture the following policy functions:

$$p_{1i}^* = \varphi_1 E(u|s_{1i}) \quad p_{2i}^* = -\varphi_2 E(u|s_{2i}) \quad x = \phi_x u, \quad (\text{C.21})$$

where firms update their beliefs according to (2.21). The conjectured solution is motivated by the fact that the solution must be linear in a linearized model, and firms' optimal actions are linear functions of their perceived state of the world.

The policy functions must satisfy firms' price-setting rules (B.3), the belief updating rule (2.21), and the monetary policy rule (2.22). Therefore, the solution must satisfy the following equations that characterize firms' responsiveness to shocks:

$$\varphi_1 = \frac{2(1-\theta_2)(1-\omega)K_2}{(1-\theta_2)(1-\omega)K_2 + (1-\theta_1)\omega K_1}, \quad (\text{C.22})$$

$$\varphi_2 = \frac{2(1-\theta_1)\omega K_1}{(1-\theta_2)(1-\omega)K_2 + (1-\theta_1)\omega K_1}, \quad (\text{C.23})$$

and $\phi_x = \frac{(1-\theta_2)(1-\omega)K_2 - (1-\theta_1)\omega K_1}{(1-\theta_2)(1-\omega)K_2 + (1-\theta_1)\omega K_1} - \frac{(1-\theta_1)(1-\theta_2)(1-\omega) - (1-\theta_1)(1-\theta_2)\omega}{(1-\theta_2)(1-\omega)K_2 + (1-\theta_1)\omega K_1} K_1 K_2$. Equations (C.21) and (B.4) imply that $p_{ki}^\diamond = \varphi_k u$ for $k = 1, 2$.

With policy functions and the information flow constraint (2.13), the optimal amount of information (2.19) can be rewritten as:

$$\kappa_{ki}^* = \arg \min_{\kappa_{ki}} \frac{(\epsilon - 1)}{2} \varphi_k^2 2^{-2\kappa_{ki}} \sigma_u^2 + \lambda f(\kappa_{ki}). \quad (\text{C.24})$$

In the baseline model, following Paciello and Wiederholt (2014), we assume an exponential cost function $f(\kappa_{ki}) = 2^{2\kappa_{ki}}$ because of its analytical convenience. Our findings are robust to alternative cost functions.

The optimal attention choice is associated with the following amount of attention (κ_{ki}):

$$\kappa_{ki} = \max \left(\frac{1}{2} \log_2 \left(\frac{(\epsilon - 1) \varphi_k^2 \sigma_u^2}{2\lambda} \right)^{\frac{1}{2}}, 0 \right) \quad (\text{C.25})$$

Note that the optimal amount of information is equal across price-resetting firms within

a sector. That is, $\kappa_{ki} = \kappa_k$. Accordingly, the optimal conditional uncertainty and Kalman gain can be written as :

$$\hat{\sigma}_{u,ki}^2 = \min \left\{ \left(\frac{2\lambda\sigma_u^2}{(\epsilon-1)\varphi_k^2} \right)^{\frac{1}{2}}, \sigma_u^2 \right\}, \quad (\text{C.26})$$

$$K_k = \max \left\{ 1 - \left(\frac{2\lambda}{(\epsilon-1)\varphi_k^2\sigma_u^2} \right)^{\frac{1}{2}}, 0 \right\}. \quad (\text{C.27})$$

Note that the amount of attention κ_{ki} and the Kalman gains (K_k) must be non-negative.

Finally, we obtain a system of equations for φ_1 and φ_2 by substituting Equation (C.27) into Equations (C.22) and (C.23).

To understand the parameters that determine the degree of information frictions, it is useful to introduce the following definition:

$$K^s \equiv 1 - \left(\frac{2\lambda}{(\epsilon-1)\sigma_u^2} \right)^{\frac{1}{2}}. \quad (\text{C.28})$$

K^s represents the Kalman gain selected by firms in a counterfactual symmetric economy with $\theta_1 = \theta_2$ and $\omega_1 = 0.5$. The analytical expression (C.28) indicates that the degree of information frictions is proportional to the cost of information acquisition: K^s decreases as λ increases.

C.3.1 Detailed Derivations of the Policy functions

We employ the method of undetermined coefficients to determine the policy functions. We guess and verify these function forms as follows:

$$\begin{aligned} p_{1i}^* &= \phi_1 s_{1i} \\ p_1^* &= \phi_1 u \\ p_{2i}^* &= -\phi_2 s_{2i} \\ p_2^* &= -\phi_2 u \\ x &= \phi_x u. \end{aligned}$$

Given our guesses, the following equations must hold:

$$\phi_1 s_{1i} = E \left[\frac{1}{2}(1 - \theta_1)\phi_1 u - \frac{1}{2}(1 - \theta_2)\phi_2 u + \phi_x u + u | s_{1i} \right] \quad (\text{C.29})$$

$$- \phi_2 s_{2i} = E \left[\frac{1}{2}(1 - \theta_1)\phi_1 u - \frac{1}{2}(1 - \theta_2)\phi_2 u + \phi_x u - u | s_{2i} \right], \quad (\text{C.30})$$

Thus, we have:

$$\phi_1 = \frac{1}{2}(1 - \theta_1)\phi_1 K_1 - \frac{1}{2}(1 - \theta_2)\phi_2 K_1 + \phi_x K_1 + K_1 \quad (\text{C.31})$$

$$- \phi_2 = \frac{1}{2}(1 - \theta_1)\phi_1 K_2 - \frac{1}{2}(1 - \theta_2)\phi_2 K_2 + \phi_x K_2 - K_2, \quad (\text{C.32})$$

where K_1 and K_2 are Kalman gains in sector 1 and 2, respectively. We assume the central bank uses a price stabilization monetary policy:

$$\frac{1}{2}(1 - \theta_1)\omega\phi_1 u - \frac{1}{2}(1 - \theta_2)(1 - \omega)\phi_2 u = 0, \quad (\text{C.33})$$

From this equation, we obtain:

$$\phi_1 = \frac{(1 - \theta_2)(1 - \omega)}{(1 - \theta_1)\omega} \phi_2 \quad (\text{C.34})$$

Plugging equation (C.34) into (C.31) and (C.32), we get

$$\frac{(1 - \theta_2)(1 - \omega)}{(1 - \theta_1)\omega} \phi_2 = \frac{1}{2}(1 - \theta_1) \frac{(1 - \theta_2)(1 - \omega)}{(1 - \theta_1)\omega} \phi_2 K_1 - \frac{1}{2}(1 - \theta_2)\phi_2 K_1 + \phi_x K_1 + K_1, \quad (\text{C.35})$$

$$- \phi_2 = \frac{1}{2}(1 - \theta_1) \frac{(1 - \theta_2)(1 - \omega)}{(1 - \theta_1)\omega} \phi_2 K_2 - \frac{1}{2}(1 - \theta_2)\phi_2 K_2 + \phi_x K_2 + K_2 \quad (\text{C.36})$$

Combining these two equations above, we can derive

$$\phi_1 = \frac{2(1 - \theta_2)(1 - \omega)K_1 K_2}{(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1}, \quad (\text{C.37})$$

$$\phi_2 = \frac{2(1 - \theta_1)\omega K_1 K_2}{(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1}. \quad (\text{C.38})$$

Plugging the two results into either equation (C.35) or equation (C.36), we obtain

$$\phi_x = \frac{(1 - \theta_2)(1 - \omega)K_2 - (1 - \theta_1)\omega K_1}{(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1} - \frac{(1 - \theta_1)(1 - \theta_2)(1 - \omega) - (1 - \theta_1)(1 - \theta_2)\omega}{(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1} K_1 K_2 \quad (\text{C.39})$$

Substituting the values of ϕ_1 , ϕ_2 , and ϕ_x into the profit-maximizing price of firm i in sector 1, we have

$$p_{1i}^\diamond = \frac{2(1 - \theta_2)(1 - \omega)K_2}{(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1} u, \quad (\text{C.40})$$

and

$$p_{1i}^* = \frac{2(1 - \theta_2)(1 - \omega)K_2}{(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1} K_1 s_{1i}, \quad (\text{C.41})$$

This equation (C.41) verifies our initial guesses. Furthermore, we define φ_1 and φ_2 as

$$\varphi_1 = \frac{\phi_1}{K_1} = \frac{2(1 - \theta_2)(1 - \omega)K_2}{(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1}, \quad (\text{C.42})$$

$$\varphi_2 = \frac{\phi_2}{K_2} = \frac{2(1 - \theta_1)\omega K_1}{(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1}. \quad (\text{C.43})$$

Note that $\varphi_1 + \varphi_2 = 2$.

D Propositions and Lemmas and Their Proofs

D.1 Lemma D.1

The Optimal Choice of Signals To simplify the presentation of the model, we have stated that, in the model, firms choose to allocate attention to one signal, equation (2.17), out of a continuum of signals (2.9). We now establish this result.

Lemma D.1. *Consider the baseline economy with a general signal structure where firms have access to a continuum of signals of the functional form: $\mathbf{s}_{ki}^\diamond = \mathbf{M}_k^\diamond \mathbf{a} + \mathbf{e}_{ki}^\diamond$ for any $\mathbf{M}_{ki}^\diamond$ and variance-covariance matrix $\Sigma_{e,k}^\diamond$. A firm ki optimally chooses to narrow its attention to one signal: $s_{ki} = u + e_{ki}$, $e_{ki} \sim N(0, \sigma_{e,k}^2)$.*

Proof. The proof closely follows Mondria (2010) and Li and Wu (2016). Let the signal structure of agents have the following general form

$$\mathbf{s}_{ki}^\diamond = \mathbf{M}_k^\diamond \mathbf{a} + \mathbf{e}_{ki}^\diamond \quad (\text{D.1})$$

where $\mathbf{a} = (a_1, a_2)'$ is the vector of sectoral shocks with variance-covariance matrix Σ_{aa} , $\mathbf{s}_{ki}^\diamond = (s_{ki,1}, s_{ki,2})'$ is the vector of signals and $\mathbf{M}_{ki}^\diamond$ is 2×2 matrix as the weights of these shocks. The vector of observational errors $\mathbf{e}_{ki}^\diamond = (e_{1,ki}^\diamond, e_{2,ki}^\diamond)'$, which follows a 2-dimensional Gaussian distribution with variance-covariance matrix $\Sigma_{e,k}^\diamond$.

In the following steps, we will show that, given the general signal structure (D.1), the agents will optimally choose the signal structure in the form:

$$s_{ki} = u + e_{ki} \quad (\text{D.2})$$

Step 1 We first take a linear transformation of the signal structure (D.1) so that the variance-covariance matrix of the error term in the new signal is diagonal. Since the variance-covariance matrix $\Sigma_{e,k}^\diamond$ is symmetric and positive semi-definite, we can always find an orthogonal matrix B_k such that

$$\Sigma_{e,k}^\diamond = B_k \Lambda_k B_k' \quad (\text{D.3})$$

where Λ_k is a diagonal matrix and $B_k' = B_k^{-1}$. First, we transform the signal linearly as follows:

$$\tilde{\mathbf{s}}_{ki}^* = B_k^{-1} \mathbf{s}_{ki}^\diamond \quad (\text{D.4})$$

$$= B_k^{-1} \mathbf{M}_k^\diamond \mathbf{a} + B_k^{-1} \mathbf{e}_{ki}^\diamond \quad (\text{D.5})$$

The variance-covariance matrix of $B_k^{-1} \mathbf{e}_{ki}^\diamond$ is a diagonal matrix Λ_k .

Then, we construct new signals $\mathbf{s}_{ki}$ of the following form:

$$\mathbf{s}_{ki} = \mathbf{M}_k \mathbf{a} + \mathbf{e}_{ki}$$

where $\mathbf{M}_k = D B_k^{-1} \mathbf{M}_k^\diamond$, $\mathbf{e}_{ki} = D B_k^{-1} \mathbf{e}_{ki}^\diamond$, and D is a diagonal non-singular matrix. We can select D such that the weighting matrix $\mathbf{M}_k$ is normalized to be $\mathbf{M}_k = \begin{bmatrix} 1 & m_{k,1} \\ 1 & m_{k,2} \end{bmatrix}$. The variance-covariance matrix of $\mathbf{e}_{ki}$ is $\Sigma_{e,k}$.

The linear transform does not alter the mutual information. In other words, the new signals $\mathbf{s}_{ki}$ satisfies the information capacity constraint and results in the same equilibrium.

Step 2 In the second step, we proceed to show that firms choose to observe only one

signal. We first describe the information capacity of firms.

Given a maximal attention κ_k , the information constraint for firms in sector k follows

$$\frac{|\mathbf{M}_k \mathbf{\Sigma}_{aa} \mathbf{M}_k' + \mathbf{\Sigma}_{e,k}|}{|\mathbf{\Sigma}_{e,k}|} \leq 2^{2\kappa_k}, \quad (\text{D.6})$$

where $|\cdot|$ denotes the determinant of a matrix. It follows that

$$\frac{\sigma_{a1}^2 + m_{k,1}^2 \sigma_{a2}^2}{\sigma_{e1,k}^2} + \frac{\sigma_{a1}^2 + m_{k,2}^2 \sigma_{a2}^2}{\sigma_{e2,k}^2} + \frac{(m_{k,1} - m_{k,2})^2 \sigma_{a1}^2 \sigma_{a2}^2}{\sigma_{e1,k}^2 \sigma_{e2,k}^2} \leq 2^{2\kappa_k}. \quad (\text{D.7})$$

where $\sigma_{e1,k}^2$ and $\sigma_{e2,k}^2$ are diagonal elements of $\mathbf{\Sigma}_{e,k}$, and σ_{a1}^2 and σ_{a2}^2 are diagonal items of $\mathbf{\Sigma}_{aa}$. The information constraint is binding in equilibrium can be rewritten compactly in terms of $\sigma_{e1,k}^2$ and $\sigma_{e2,k}^2$ as $y_{k,1} \sigma_{e1,k}^{-2} + y_{k,2} \sigma_{e2,k}^{-2} + y_{k,3} \sigma_{e1,k}^{-2} \sigma_{e2,k}^{-2} = y_{k,4}$, where $y_{k,1}$, $y_{k,2}$, $y_{k,3}$ and $y_{k,4}$ are positive real numbers and defined as

$$\begin{aligned} y_{k,1} &= \sigma_{a1}^2 + m_{k,1}^2 \sigma_{a2}^2, \\ y_{k,2} &= \sigma_{a1}^2 + m_{k,2}^2 \sigma_{a2}^2, \\ y_{k,3} &= (m_{k,1} - m_{k,2})^2 \sigma_{a1}^2 \sigma_{a2}^2, \\ y_{k,4} &= 2^{2\kappa_k}. \end{aligned}$$

Next, we describe the optimization problem of firms. Given the information capacity κ_k , firm i in sector k minimizes the profit loss function $E((p_{ki}^* - p_{ki}^\diamond)^2 | \mathbf{s}_{ki})$ subject to the information constraint (D.6).

Assume that in equilibrium $p_{ki}^\diamond = \mathbf{\Phi}_{p_k^\diamond} \mathbf{a}$, where $\mathbf{\Phi}_{p_k^\diamond} = (\phi_{p_k^\diamond,1}, \phi_{p_k^\diamond,2})$. This can be easily verified later.²⁰ Using the information constraint (D.6) and the conjectured equilibrium

²⁰We will show that $p_{ki}^\diamond$ is actually a function of u , defined as $\frac{1}{2}(a_2 - a_1)$, as shown later.

for $p_{ki}^\diamond$, we derive the profit loss function:

$$\begin{aligned}
& E((p_{ki}^* - p_{ki}^\diamond)^2 \mid \mathbf{s}_{ki}) \\
= & \text{Var}(p_{ki}^\diamond \mid \mathbf{s}_{ki}) \\
= & \Phi_{p_k^\diamond} \Sigma_{aa} \Phi_{p_k^\diamond}' - \Phi_{p_k^\diamond} \Sigma_{aa} \mathbf{M}_k' (\mathbf{M}_k \Sigma_{aa} \mathbf{M}_k' + \Sigma_{e,k})^{-1} \mathbf{M}_k \Sigma_{aa} \Phi_{p_k^\diamond}' \\
= & \Phi_{p_k^\diamond} \Sigma_{aa} \Phi_{p_k^\diamond}' - \Phi_{p_k^\diamond} \Sigma_{aa} \mathbf{M}_k' \frac{\text{Tr}(\mathbf{M}_k \Sigma_{aa} \mathbf{M}_k' + \Sigma_{e,k}) \mathbf{1} - (\mathbf{M}_k \Sigma_{aa} \mathbf{M}_k' + \Sigma_{e,k})}{2^{2\kappa_k} \det(\Sigma_{k,e})} \mathbf{M}_k \Sigma_{aa} \Phi_{p_k^\diamond}' \\
= & x_{k,0} + x_{k,1} \sigma_{e1,k}^{-2} + x_{k,2} \sigma_{e2,k}^{-2} + x_{k,3} \sigma_{e1,k}^{-2} \sigma_{e2,k}^{-2}
\end{aligned}$$

where the first equality uses the pricing equation $p_{ki}^* = E(p_{ki}^\diamond \mid \mathbf{s}_{ki})$, the second equality uses the equilibrium solution, the third equality uses equation (D.6) and inverse operation of a matrix, and the fourth equality is obtained with some algebra. Tr means the trace of a matrix and $\mathbf{1}$ is 2×2 identity matrix. Coefficients $x_{k,0}$, $x_{k,1}$, $x_{k,2}$ and $x_{k,3}$ are all positive real numbers:

$$\begin{aligned}
x_{k,0} &= \phi_{1,p_k^\diamond}^2 \sigma_{a1}^2 + \phi_{2,p_k^\diamond}^2 \sigma_{a2}^2 \\
x_{k,1} &= 2^{-2\kappa_k} \left(\phi_{1,p_k^\diamond}^2 \sigma_{a1}^2 + m_{k,1} \phi_{2,p_k^\diamond}^2 \sigma_{a2}^2 \right)^2 \\
x_{k,2} &= 2^{-2\kappa_k} \left(\phi_{1,p_k^\diamond}^2 \sigma_{a1}^2 + m_{k,2} \phi_{2,p_k^\diamond}^2 \sigma_{a2}^2 \right)^2 \\
x_{k,3} &= 2^{-2\kappa_k} (m_{k,1} - m_{k,2})^2 \sigma_{a1}^2 \sigma_{a2}^2 \left(\phi_{1,p_k^\diamond}^2 \sigma_{a2}^2 + \phi_{2,p_k^\diamond}^2 \sigma_{a1}^2 \right).
\end{aligned}$$

Therefore, firms' attention allocation problem can be reformulated as

$$\begin{aligned}
& \min_{\sigma_{e1,k}^2, \sigma_{e2,k}^2} x_{k,0} + x_{k,1} \sigma_{e1,k}^{-2} + x_{k,2} \sigma_{e2,k}^{-2} + x_{k,3} \sigma_{e1,k}^{-2} \sigma_{e2,k}^{-2} \\
& \text{s.t. } y_{k,4} = y_{k,1} \sigma_{e1,k}^{-2} + y_{k,2} \sigma_{e2,k}^{-2} + y_{k,3} \sigma_{e1,k}^{-2} \sigma_{e2,k}^{-2}
\end{aligned}$$

Substituting constraints into the objective function, we obtain

$$\min_{\sigma_{e1,k}^2 \in \left[0, \frac{y_{k,4}}{y_{k,1}}\right]} x_{k,0} + x_{k,1} \sigma_{e1,k}^{-2} + \left(x_{k,1} + x_{k,3} \sigma_{e1,k}^{-2} \right) \frac{y_{k,4} - y_{k,1} \sigma_{e1,k}^{-2}}{y_{k,2} + y_{k,3} \sigma_{e1,k}^{-2}}$$

The second-order condition is

$$\begin{aligned}
& \frac{(y_1 y_2 + y_3 y_4)}{(y_2 + y_3 z_1)^3} (x_2 y_3 - y_2 x_3) \\
& = -2^{-2\kappa_k} (m_{k,1} - m_{k,2})^2 \sigma_{a1}^4 \sigma_{a2}^4 \left(\phi_{1,p_k^\diamond} m_{k,2} - \phi_{2,p_k^\diamond} \right)^2 \\
& < 0.
\end{aligned}$$

The objective function is strictly concave over a compact set $\left[0, \frac{y_{k,4}}{y_{k,1}}\right]$. There exist only corner solutions, which implies that firms choose to observe only one signal. The firms choose to observe either signal $s_{ki,1}$ (i.e. $\sigma_{e1,k}^{-2} = y_{k,4}/y_{k,1}, \sigma_{e2,k}^{-2} = 0$) or signal $s_{ki,2}$ (i.e. $\sigma_{e1,k}^{-2} = 0, \sigma_{e2,k}^{-2} = y_{k,4}/y_{k,2}$). Without loss of the generality, we assume firms choose to observe:

$$\begin{aligned}
s_{ki} &= a_1 + m_{k,1} a_2 + e_{ki} \\
&= \tilde{M}_k \mathbf{a} + e_{ki}
\end{aligned}$$

where $\tilde{M}_k = (1, m_{k,1})$.

Step 3: In the third step, we characterize the optimal information structure that firms choose to observe. Specifically, we solve for the optimal weights $\tilde{M}_k$ and variance σ_e^2 for firms in sector k as follows:

$$\begin{aligned}
\min_{m_{k,1}, \sigma_e^2} & \Phi_{p_k^\diamond \Sigma_{aa}} \Phi_{p_k^\diamond}' - \frac{\left(\Phi_{p_k^\diamond \Sigma_{aa}} \tilde{M}_k' \right)^2}{\tilde{M}_k \Sigma_{aa} \tilde{M}_k' + \sigma_e^2} \\
\text{s.t.} & \frac{\tilde{M}_k \Sigma_{aa} \tilde{M}_k' + \sigma_e^2}{\sigma_e^2} = 2^{2\kappa_k}
\end{aligned}$$

The solution is

$$\begin{aligned}
m_{k,1} &= \frac{\phi_{2,p_k^\diamond}}{\phi_{1,p_k^\diamond}} \\
\sigma_e^2 &= \frac{1}{2^{2\kappa_k} - 1} \tilde{M}_k \Sigma_{aa} \tilde{M}_k'.
\end{aligned}$$

Thus $\tilde{M}_k = \left(1, \frac{\phi_{2,p_k^\diamond}}{\phi_{1,p_k^\diamond}}\right)$. Take linear transform of the signal, it turns out that the optimal signal for firms in sector k is a signal about $p_{ki}^\diamond = \left(\phi_{p_k^\diamond,1}, \phi_{p_k^\diamond,2}\right) \mathbf{a}$. With a little abuse of

notation, it is

$$s_{ki} = p_{ki}^\diamond + e_{ki}.$$

Step 4 In the last step, we show that $p_{ki}^\diamond$ is a linear function of u and the optimal signal for firms is a signal about u . Remember that $u \equiv \frac{1}{2}(a_2 - a_1)$. Given the information structure (D.8), and the systems of the equations

$$\begin{aligned} p_{ki}^* &= E(p_{ki}^\diamond | s_{ki}) \\ p_{1i}^\diamond &= p + x + u \\ p_{2i}^\diamond &= p + x - u \\ 0 &= \omega p_1 + (1 - \omega)p_2 \\ p_k &= (1 - \theta_k) \int p_{ki}^* di, \end{aligned}$$

$p_{ki}^\diamond$ is linear function of u . Thus, by the linear transformation, the optimal signal is a signal about u equivalently. With a little abuse of notation, it is

$$s_{ki} = u + e_{ki}.$$

□

D.2 Proof of Proposition 1

Proof. For a re-optimizing firm, its optimal attention problem leads to

$$f'(\kappa = \kappa_1) = \frac{(\epsilon - 1)\sigma_u^2}{\lambda} \varphi_1^2 2^{-2\kappa_1} \ln 2, \quad (\text{D.8})$$

$$f'(\kappa = \kappa_2) = \frac{(\epsilon - 1)\sigma_u^2}{\lambda} \varphi_2^2 2^{-2\kappa_2} \ln 2 \quad (\text{D.9})$$

In addition, we already have

$$\varphi_1 = \frac{2(1 - \theta_2)(1 - \omega)K_2}{(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1} \quad (\text{D.10})$$

$$\varphi_2 = \frac{2(1 - \theta_1)\omega K_1}{(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1}, \quad (\text{D.11})$$

where

$$\begin{aligned} K_1 &= 1 - 2^{-2\kappa_1} \\ K_2 &= 1 - 2^{-2\kappa_2}. \end{aligned}$$

Let us focus on sector 1 without loss of generality. First, note that if the information friction is exogenous, it is straightforward to show that $\frac{\partial \varphi_1}{\partial \theta_1} > 0$. However, due to rational inattention, the information choice and the policy function are jointly determined. In what follows, we demonstrate that this property holds under certain conditions. Moreover, we establish this result using a proof by contradiction.

In the following, to simplify notation, we use $f'(\kappa_1)$ to represent $f'(\kappa = \kappa_1)$, and $f''(\kappa_1)$ to represent $f''(\kappa = \kappa_1)$, respectively. Take derivative with respect to θ_1 for (D.8):

$$\frac{(\epsilon - 1)\sigma_u^2}{\lambda} 2^{2\kappa_1} \left(f''(\kappa_1) + 2\ln 2 f'(\kappa_1) \right) \frac{\partial \kappa_1}{\partial \theta_1} = 2\varphi_1 \frac{\partial \varphi_1}{\partial \theta_1}$$

We have assumed that $f'(\kappa) > 0$. In addition, the policy function φ_1 is positive and $\epsilon > 1$ is a standard calibration. Thus, if and only if $\frac{f''(\kappa)}{f'(\kappa)} > -2\ln 2$, we have $f''(\kappa_1) + 2\ln 2 f'(\kappa_1) > 0$, and $\frac{\partial \kappa_1}{\partial \theta_1}$ has the same sign as $\frac{\partial \varphi_1}{\partial \theta_1}$. Similarly, $\frac{\partial \kappa_2}{\partial \theta_1}$ has the same sign as $\frac{\partial \varphi_2}{\partial \theta_1}$ if and only if $\frac{f''(\kappa)}{f'(\kappa)} > -2\ln 2$.

Take the derivative of φ_1 with respect to θ_1 with equation (D.10), its numerator is (ignoring the denominator because we care about its sign and the denominator must be positive):

$$\begin{aligned} & 2(1 - \theta_2)(1 - \omega) \frac{\partial K_2}{\partial \theta_1} [(1 - \theta_2)(1 - \omega)K_2 + (1 - \theta_1)\omega K_1] \\ & - \left[(1 - \theta_2)(1 - \omega) \frac{\partial K_2}{\partial \theta_1} + (1 - \theta_1)\omega \frac{\partial K_1}{\partial \theta_1} - \omega K_1 \right] 2(1 - \theta_2)(1 - \omega)K_2 \\ & = 2(1 - \theta_2)(1 - \omega)\omega \left[\frac{\partial K_2}{\partial \theta_1} (1 - \theta_1)K_1 - (1 - \theta_1) \frac{\partial K_1}{\partial \theta_1} K_2 + K_1 K_2 \right] \end{aligned} \quad (\text{D.12})$$

We now demonstrate $\frac{\partial K_1}{\partial \theta_1} > 0$, using the proof by contradiction. Otherwise, suppose $\frac{\partial K_1}{\partial \theta_1} < 0$, then $\frac{\partial \varphi_1}{\partial \theta_1} < 0$. Thus $\frac{\partial \varphi_2}{\partial \theta_1} > 0$ as a result of $\varphi_1 + \varphi_2 = 2$, and hence $\frac{\partial K_2}{\partial \theta_1} > 0$. However, $\frac{\partial K_2}{\partial \theta_1} > 0$ and $\frac{\partial K_1}{\partial \theta_1} < 0$ would imply that $\frac{\partial \varphi_1}{\partial \theta_1} > 0$ as can be seen with equation (D.12), which contradicts $\frac{\partial \varphi_1}{\partial \theta_1} < 0$. Therefore, $\frac{\partial K_1}{\partial \theta_1} > 0$ and $\frac{\partial \varphi_1}{\partial \theta_1} > 0$. □

D.3 Proof of Lemma 3.1

Proof. We first characterize the equilibrium under the Ramsey optimal monetary policy and then show that the equilibrium can be implemented by a simple price stabilization rule in the form $\omega_1^* p_1 + \omega_2^* p_2 = 0$.

As defined in section 2.4, the Ramsey policy optimally targets the aggregate nominal demand, given by $p + x = \phi u$, by choosing the optimal ϕ to minimize the unconditional welfare loss.

Step 1: Given ϕ , all other endogenous variables are determined. Specifically,

$$p_{1i}^* = (\phi + 1)\mathbb{E}[u|I_{1i}]$$

$$p_{2i}^* = (\phi - 1)\mathbb{E}[u|I_{2i}]$$

The Kalman gain is

$$K_k = 1 - \left(\frac{2\lambda}{(\epsilon - 1)\varphi_k^2 \sigma_u^2} \right)^{\frac{1}{2}} = 1 - \frac{1 - K^S}{\varphi_k} \quad (\text{D.13})$$

where

$$\varphi_1 = \phi + 1, \quad \varphi_2 = 1 - \phi. \quad (\text{D.14})$$

And

$$K^S \equiv 1 - \left(\frac{2\lambda}{(\epsilon - 1)\sigma_u^2} \right)^{\frac{1}{2}}. \quad (\text{D.15})$$

Here we focus on the parameter region for λ that $K_k > 0$. Note that

$$(\phi + 1)K_1 = \varphi_1 \left(1 - \frac{1 - K^S}{\varphi_1} \right) = \varphi_1 - 1 + K^S = \phi + K^S \quad (\text{D.16})$$

$$(\phi - 1)K_2 = -\varphi_2 \left(1 - \frac{1 - K^S}{\varphi_2} \right) = -\varphi_2 + 1 - K^S = \phi - K^S \quad (\text{D.17})$$

The total welfare loss function can be decomposed into four parts: (1) price dispersion within sectors due to sticky prices $\mathbb{L}_{Calvo}$, (2) price dispersion within the sector due to dispersed information $\mathbb{L}_{DS}$, (3) output gap $\mathbb{L}_x$, and (4) relative price gap across sectors

$\mathbb{L}_{rp}$. The total welfare loss function is

$$\mathbb{L}_{total} = \mathbb{L}_{Calvo} + \mathbb{L}_{DS} + \mathbb{L}_x + \mathbb{L}_{rp}$$

where

$$\begin{aligned}\mathbb{L}_{Calvo} &= \frac{\varepsilon}{2} \left[(1 - \theta_1)\theta_1(\phi + 1)^2 K_1^2 \sigma_u^2 + (1 - \theta_2)\theta_2(\phi - 1)^2 K_2^2 \sigma_u^2 \right] \\ \mathbb{L}_{DS} &= \frac{\varepsilon}{2} \left[(1 - \theta_1)(\phi + 1)^2 K_1 \sigma_{e,1}^2 + (1 - \theta_2)(\phi - 1)^2 K_2^2 \sigma_{e,2}^2 \right] \\ \mathbb{L}_x &= \sigma_u^2 \left[\left(1 - \frac{1 - \theta_1}{2} K_1 - \frac{1 - \theta_2}{2} K_2 \right) \phi + \frac{1 - \theta_2}{2} K_2 - \frac{1 - \theta_1}{2} K_1 \right]^2 \\ \mathbb{L}_{rp} &= \frac{\eta}{4} \sigma_u^2 \left\{ [(1 - \theta_1)K_1 - (1 - \theta_2)K_2] \phi + (1 - \theta_1)K_1 + (1 - \theta_2)K_2 - 2 \right\}^2\end{aligned}$$

Step 2: Second, we determine the optimal ϕ that minimizes the total welfare loss. To do this, we calculate the first-order derivative of the total welfare loss, $\mathbb{L}_{total}$, with respect to ϕ and set it equal to zero:

$$\frac{\partial \mathbb{L}_{total}}{\partial \phi} = \frac{\partial \mathbb{L}_{Calvo}}{\partial \phi} + \frac{\partial \mathbb{L}_{DS}}{\partial \phi} + \frac{\partial \mathbb{L}_x}{\partial \phi} + \frac{\partial \mathbb{L}_{rp}}{\partial \phi} = 0$$

Let ϕ^* be the solution to the equation [D.4](#), we have

$$\phi^* = \frac{D_2}{D_1}$$

where D_1 and D_2 are defined as follows:

$$D_1 \equiv \varepsilon \left[(1 - \theta_1)\theta_1 + (1 - \theta_2)\theta_2 \right] + \frac{1}{2}(\theta_1 + \theta_2)^2 + \frac{\eta}{2}(\theta_2 - \theta_1)^2 \quad (\text{D.18})$$

$$-D_2 \equiv \varepsilon \left[(1 - \theta_1) \left(\frac{1 - K^S}{2} + \theta_1 K^S \right) - (1 - \theta_2) \left(\frac{1 - K^S}{2} + \theta_2 K^S \right) \right] \quad (\text{D.19})$$

$$+ \frac{1}{2}(\theta_1 + \theta_2)(\theta_1 - \theta_2)K^S \quad (\text{D.20})$$

$$+ \eta \left[\left(1 - \frac{\theta_1 + \theta_2}{2} \right) K^S - 1 \right] (\theta_2 - \theta_1) \quad (\text{D.21})$$

Step 3: Third, we derive the optimal ω_1^* and ω_2^* to implement the same equilibrium allocation as that under Ramsey optimal policy. The price index policy is specified as follows:

$$\omega_1 p_1 + \omega_2 p_2 = 0 \quad (\text{D.22})$$

Let p_1 and p_2 be the allocation under the Ramsey optimal policy; we can always find a ω_1^* and ω_2^* such that (D.22) holds. Under the optimal policy, $p_1 = (1 - \theta_1)(\phi + 1)K_1 u$ and $p_2 = (1 - \theta_2)(\phi - 1)K_2 u$,

Then, we have

$$\omega_1^*(1 - \theta_1)(\phi + 1)K_1 u + \omega_2^*(1 - \theta_2)(\phi - 1)K_2 u = 0$$

Use the conditions (D.13) and (D.55), We can derive that

$$\frac{\omega_1^*}{\omega_2^*} = \frac{(1 - \theta_2)(K^S - \phi)}{(1 - \theta_1)(K^S + \phi)} = \frac{(1 - \theta_2)(D_1 K^S - D_2)}{(1 - \theta_1)(D_1 K^S + D_2)}$$

The price weight ω_1^* and ω_2^* are not unique. Without the loss of generality, we choose

$$\omega_1^* = (1 - \theta_2)(D_1 K^S - D_2)$$

$$\omega_2^* = (1 - \theta_1)(D_1 K^S + D_2)$$

Substitute D_1 and D_2 using (D.18) and (D.21). Then, we have

$$\begin{aligned} \omega_1^* &= (1 - \theta_2) \left\{ \varepsilon \left[(\theta_2 - \theta_1) \frac{1 - K^S}{2} + (1 - \theta_1) 2\theta_1 K^S \right] + (\theta_1 + \theta_2) \theta_1 K^S + \eta \left[(1 - \theta_1) K^S - 1 \right] (\theta_2 - \theta_1) \right\} \\ \omega_2^* &= (1 - \theta_1) \left\{ \varepsilon \left[(\theta_1 - \theta_2) \frac{1 - K^S}{2} + (1 - \theta_2) 2\theta_2 K^S \right] + (\theta_1 + \theta_2) \theta_2 K^S + \eta \left[(1 - \theta_2) K^S - 1 \right] (\theta_1 - \theta_2) \right\}. \end{aligned}$$

□

D.4 Proof of Proposition 2

Proof. We follow the same procedure in D.3 to derive the optimal price index to replicate the allocations under Ramsey optimal policy. Then we show that $\omega_1^* < \omega_2^*$ if $\theta_1 < \theta_2$.

Step 1 : Given ϕ , all other endogenous variables are determined and can be derived. Specifically,

$$p_{1i}^* = (\phi + 1)\mathbb{E}[u|I_{1i}]$$

$$p_{2i}^* = (\phi - 1)\mathbb{E}[u|I_{2i}]$$

Given ϕ , we can derive solutions for all endogenous variables. Then, we substitute the solutions into the total welfare loss function.

The total welfare loss function is shown as follows

$$\mathbb{L}_{total} = \mathbb{L}_{Calvo} + \mathbb{L}_{DS} + \mathbb{L}_x + \mathbb{L}_{rp}$$

where

$$\mathbb{L}_{Calvo} = \frac{\epsilon}{2} \left\{ (1 - \theta_1)\theta_1(\phi + 1)^2 K_1^2 \sigma_u^2 + (1 - \theta_2)\theta_2(\phi - 1)^2 K_2^2 \sigma_u^2 \right\}$$

$$\mathbb{L}_{DS} = \frac{\epsilon}{2} \left\{ (1 - \theta_1)(\phi + 1)^2 K_1^2 \sigma_{e,1}^2 + (1 - \theta_2)(\phi - 1)^2 K_2^2 \sigma_{e,2}^2 \right\}$$

$$\mathbb{L}_x = \left(\phi - \frac{1 - \theta_1}{2} K_1 \phi - \frac{1 - \theta_1}{2} K_1 - \frac{1 - \theta_2}{2} K_2 \phi + \frac{1 - \theta_2}{2} K_2 \right)^2 \sigma_u^2$$

$$\begin{aligned} \mathbb{L}_{rp} &= \frac{\eta}{4} \mathbb{E} \left\{ (p_1 - p_2 - 2u)^2 \right\} \\ &= \frac{\eta}{4} \left\{ [(1 - \theta_1) K_1 - (1 - \theta_2) K_2] \phi + (1 - \theta_1) K_1 + (1 - \theta_2) K_2 - 2 \right\}^2 \sigma_u^2 \end{aligned}$$

Step 2 Second, we determine the optimal ϕ that minimizes the total welfare loss. To do this, we calculate the first-order derivative of the total welfare loss, $\mathbb{L}_{total}$, with respect to ϕ and set it equal to zero:

$$\frac{\partial \mathbb{L}_{total}}{\partial \phi} = \frac{\partial \mathbb{L}_{Calvo}}{\partial \phi} + \frac{\partial \mathbb{L}_{DS}}{\partial \phi} + \frac{\partial \mathbb{L}_x}{\partial \phi} + \frac{\partial \mathbb{L}_{rp}}{\partial \phi} = 0$$

The optimal ϕ^* is given by:

$$\phi^* = \frac{D_2}{D_1} \equiv \frac{-(2A_2 + \frac{\eta}{2}B_2 + \epsilon C_2)}{2A_1 + \frac{\eta}{2}B_1 + \epsilon C_1}$$

where

$$C_1 = [1 - (1 - \theta_1) K_1] (1 - \theta_1) K_1 + [1 - (1 - \theta_2) K_2] (1 - \theta_2) K_2$$

$$C_2 = [1 - (1 - \theta_1) K_1] (1 - \theta_1) K_1 - [1 - (1 - \theta_2) K_2] (1 - \theta_2) K_2$$

$$A_1 = \left(1 - \frac{1 - \theta_1}{2} K_1 - \frac{1 - \theta_2}{2} K_2\right)^2$$

$$A_2 = \left(\frac{1 - \theta_2}{2} K_2 - \frac{1 - \theta_1}{2} K_1\right) \left(1 - \frac{1 - \theta_1}{2} K_1 - \frac{1 - \theta_2}{2} K_2\right)$$

$$B_1 = [(1 - \theta_1) K_1 - (1 - \theta_2) K_2]^2$$

$$B_2 = [(1 - \theta_1) K_1 + (1 - \theta_2) K_2 - 2] [(1 - \theta_1) K_1 - (1 - \theta_2) K_2]$$

Step 3 Third, we find the optimal ω_1^* and ω_2^* to implement the allocation under the optimal monetary policy. The price index policy is specified as follows:

$$\omega_1 p_1 + \omega_2 p_2 = 0 \quad (\text{D.23})$$

Let p_1 and p_2 be the allocation under the Ramsey optimal policy; we can always find a ω_1^* and ω_2^* such that (D.23) holds. Under the optimal policy, $p_1 = (1 - \theta_1)(\phi^* + 1)K_1 u$ and $p_2 = (1 - \theta_2)(\phi^* - 1)K_2 u$. Then, we have

$$\omega_1^* (1 - \theta_1)(\phi + 1)K_1 u + \omega_2^* (1 - \theta_2)(\phi - 1)K_2 u = 0$$

Therefore,

$$\frac{\omega_1^*}{\omega_2^*} = -\frac{(1 - \theta_2)(\phi^* - 1)K_2}{(1 - \theta_1)(\phi^* + 1)K_1} = \frac{(D_1 - D_2)(1 - \theta_2)K_2}{(D_1 + D_2)(1 - \theta_1)K_1}$$

Without loss of generality, we choose

$$\omega_1^* = (D_1 - D_2)(1 - \theta_2)K_2$$

$$\omega_2^* = (D_1 + D_2)(1 - \theta_1)K_1$$

If $K_1 = K_2 = K$, then we have the optimal price index:

$$\begin{aligned}\omega_1^* &= 2\epsilon(1 - \theta_1)K^2 [1 - K + \theta_1 K] (1 - \theta_2) \\ &\quad + 2 \left(1 - \frac{1 - \theta_1}{2} K - \frac{1 - \theta_2}{2} K \right) [1 - (1 - \theta_1)K] (1 - \theta_2)K \\ &\quad + \frac{\eta}{2} [(1 - \theta_1)K - (1 - \theta_2)K] [2(1 - \theta_1)K - 2] (1 - \theta_2)K\end{aligned}$$

$$\begin{aligned}\omega_2^* &= 2\epsilon(1 - \theta_2)K^2 [1 - K + \theta_2 K] (1 - \theta_1) \\ &\quad + 2 \left(1 - \frac{1 - \theta_1}{2} K - \frac{1 - \theta_2}{2} K \right) [1 - (1 - \theta_2)K] (1 - \theta_1)K \\ &\quad + \frac{\eta}{2} [(1 - \theta_1)K - (1 - \theta_2)K] [-2(1 - \theta_2)K + 2] (1 - \theta_1)K.\end{aligned}$$

Step 4 Finally, once the expression for the optimal price index is derived, it follows directly that if $\theta_1 < \theta_2$, then $\omega_1^* < \omega_2^*$. This completes the proof. $\square$

D.5 Proof of Proposition 3

Proof. From D.3, we have

$$\begin{aligned}\omega_1^* &= (1 - \theta_2) \left\{ \epsilon \left[(\theta_2 - \theta_1) \frac{1 - K^S}{2} + (1 - \theta_1) 2\theta_1 K^S \right] + (\theta_1 + \theta_2) \theta_1 K^S + \eta \left[(1 - \theta_1) K^S - 1 \right] (\theta_2 - \theta_1) \right\} \\ \omega_2^* &= (1 - \theta_1) \left\{ \epsilon \left[(\theta_1 - \theta_2) \frac{1 - K^S}{2} + (1 - \theta_2) 2\theta_2 K^S \right] + (\theta_1 + \theta_2) \theta_2 K^S + \eta \left[(1 - \theta_2) K^S - 1 \right] (\theta_1 - \theta_2) \right\}\end{aligned}$$

The difference between ω_1^* and ω_2^* :

$$\begin{aligned}
\omega_1^* - \omega_2^* &= \varepsilon \left[2(1 - \theta_1)(1 - \theta_2)(\theta_1 - \theta_2)K^S - (2 - \theta_1 - \theta_2)(\theta_1 - \theta_2)\frac{1 - K^S}{2} \right] \\
&\quad + (\theta_1 - \theta_2)(\theta_1 + \theta_2)K^S \\
&\quad + \eta(\theta_1 - \theta_2) \left[- (1 - \theta_2)(1 - \theta_1)2K^S + (2 - \theta_1 - \theta_2) \right] \\
&= \varepsilon(\theta_1 - \theta_2) \left[2(1 - \theta_1)(1 - \theta_2)K^S - (2 - \theta_1 - \theta_2)\frac{1 - K^S}{2} \right] \\
&\quad + (\theta_1 - \theta_2)(\theta_1 + \theta_2)K^S \\
&\quad + \eta(\theta_1 - \theta_2) \left[- (1 - \theta_2)(1 - \theta_1)2K^S + (2 - \theta_1 - \theta_2) \right]
\end{aligned}$$

Then, it is straightforward to show that given $\theta_1 < \theta_2$,

$$\omega_1^* > \omega_2^* \text{ if and only if } K^S < \bar{K},$$

where

$$\bar{K} = \frac{(\frac{\varepsilon}{2} - \eta)(2 - \theta_1 - \theta_2)}{\frac{\varepsilon}{2}(2 - \theta_1 - \theta_2) + 2(\varepsilon - \eta)(1 - \theta_1)(1 - \theta_2) + (\theta_1 + \theta_2)}.$$

□

D.6 Proof of Lemmma 3.2

Proof. We can repeat the procedure in (D.4) to derive the optimal index for each component in the welfare loss. If information frictions are exogenous, the optimal index is given by:

$$\omega_{1,Calvo}^* = 2\varepsilon(1 - \theta_1)K^2 [\theta_1 K] (1 - \theta_2) \quad (\text{D.24})$$

$$\omega_{1,DispB}^* = 2\varepsilon(1 - \theta_1)K^2 [1 - K] (1 - \theta_2) \quad (\text{D.25})$$

$$\omega_{1,OutputGap}^* = 2 \left(1 - \frac{1 - \theta_1}{2}K - \frac{1 - \theta_2}{2}K \right) [1 - (1 - \theta_1)K] (1 - \theta_2)K \quad (\text{D.26})$$

$$\omega_{1,RPGap}^* = \frac{\eta}{2} [(1 - \theta_1)K - (1 - \theta_2)K] [2(1 - \theta_1)K - 2] (1 - \theta_2)K \quad (\text{D.27})$$

$$\omega_{2,Calvo}^* = 2\epsilon(1 - \theta_2)K^2 [\theta_2 K] (1 - \theta_1) \quad (D.28)$$

$$\omega_{2,DispB}^* = 2\epsilon(1 - \theta_2)K^2 [1 - K] (1 - \theta_1) \quad (D.29)$$

$$\omega_{2,OutputGap}^* = 2 \left(1 - \frac{1 - \theta_1}{2}K - \frac{1 - \theta_2}{2}K \right) [1 - (1 - \theta_2)K] (1 - \theta_1)K \quad (D.30)$$

$$\omega_{2,RPGap}^* = \frac{\eta}{2} [(1 - \theta_1)K - (1 - \theta_2)K] [-2(1 - \theta_2)K + 2] (1 - \theta_1)K \quad (D.31)$$

From above equations, it is straightforward to show that if $\theta_1 < \theta_2$ and $K > 0$, then

- $\omega_{1,Calvo}^* < \omega_{2,Calvo}^*$, $\omega_{1,OutputGap}^* < \omega_{2,OutputGap}^*$, and $\omega_{1,RPGap}^* < \omega_{2,RPGap}^*$;
- $\omega_{1,DispB}^* = \omega_{2,DispB}^*$.

If the information frictions are endogenous, the optimal index is given by:

$$\omega_{1,Calvo}^* = (1 - \theta_2)\epsilon(1 - \theta_1)2\theta_1 K^S \quad (D.32)$$

$$\omega_{1,DispB}^* = (1 - \theta_2)\epsilon \left[(\theta_2 - \theta_1) \frac{1 - K^S}{2} \right] \quad (D.33)$$

$$\omega_{1,OutputGap}^* = (1 - \theta_2)(\theta_1 + \theta_2)\theta_1 K^S \quad (D.34)$$

$$\omega_{1,RPGap}^* = (1 - \theta_2)\eta \left[(1 - \theta_1)K^S - 1 \right] (\theta_2 - \theta_1) \quad (D.35)$$

$$\omega_{2,Calvo}^* = (1 - \theta_1)\epsilon(1 - \theta_2)2\theta_2 K^S \quad (D.36)$$

$$\omega_{2,DispB}^* = (1 - \theta_1)\epsilon \left[(\theta_1 - \theta_2) \frac{1 - K^S}{2} \right] \quad (D.37)$$

$$\omega_{2,OutputGap}^* = (1 - \theta_1)(\theta_1 + \theta_2)\theta_2 K^S \quad (D.38)$$

$$\omega_{2,RPGap}^* = (1 - \theta_1)\eta \left[(1 - \theta_2)K^S - 1 \right] (\theta_1 - \theta_2) \quad (D.39)$$

Therefore, it is straightforward to show that if $\theta_1 < \theta_2$, then

- $\omega_{1,Calvo}^* < \omega_{2,Calvo}^*$, $\omega_{1,OutputGap}^* < \omega_{2,OutputGap}^*$, and $\omega_{1,RPGap}^* < \omega_{2,RPGap}^*$;
- $\omega_{1,DispB}^* > \omega_{2,DispB}^*$.

Moreover, from the equation below, we can see the difference between $\omega_{1,DispB}^*$ and $\omega_{2,DispB}^*$ is amplified when K^S is smaller.

$$\omega_{1,DispB}^* - \omega_{2,DispB}^* = \epsilon(\theta_2 - \theta_1) \frac{1 - K^S}{2} (2 - \theta_2 - \theta_1) \quad (D.40)$$

□

D.7 Proof of Corollary 4

Proof. From D.5, the difference between ω_1^* and ω_2^* is:

$$\begin{aligned}
\omega_1^* - \omega_2^* &= \varepsilon \left[2(1 - \theta_1)(1 - \theta_2)(\theta_1 - \theta_2)K^S - (2 - \theta_1 - \theta_2)(\theta_1 - \theta_2)\frac{1 - K^S}{2} \right] \\
&\quad + (\theta_1 - \theta_2)(\theta_1 + \theta_2)K^S \\
&\quad + \eta(\theta_1 - \theta_2) \left[- (1 - \theta_2)(1 - \theta_1)2K^S + (2 - \theta_1 - \theta_2) \right] \\
&= \varepsilon(\theta_1 - \theta_2) \left[2(1 - \theta_1)(1 - \theta_2)K^S - (2 - \theta_1 - \theta_2)\frac{1 - K^S}{2} \right] \\
&\quad + (\theta_1 - \theta_2)(\theta_1 + \theta_2)K^S \\
&\quad + \eta(\theta_1 - \theta_2) \left[- (1 - \theta_2)(1 - \theta_1)2K^S + (2 - \theta_1 - \theta_2) \right]
\end{aligned}$$

Then, it is straightforward to show that $\forall \theta_1, \theta_2, \omega_1^* = \omega_2^*$ if $K^S = \bar{K}$, where

$$\bar{K} = \frac{(\frac{\varepsilon}{2} - \eta)(2 - \theta_1 - \theta_2)}{\frac{\varepsilon}{2}(2 - \theta_1 - \theta_2) + 2(\varepsilon - \eta)(1 - \theta_1)(1 - \theta_2) + (\theta_1 + \theta_2)}.$$

□

D.8 Proof of Proposition 5

Proof. In the analysis below, we focus on the case where the solution is interior. The optimal attention choice problem leads to the following equations:

$$K_1 = 1 - \frac{1 - K^S}{\varphi_1} \quad (\text{D.41})$$

$$K_2 = 1 - \frac{1 - K^S}{\varphi_2}. \quad (\text{D.42})$$

where $K^S = 1 - \left(\frac{2\lambda}{(\varepsilon - 1)\sigma_u^2} \right)^{\frac{1}{2}}$.

The price stabilization policy indicates that

$$\omega(1 - \theta_1)\varphi_1 K_1 - (1 - \omega)(1 - \theta_2)\varphi_2 K_2 = 0 \quad (\text{D.43})$$

where $\omega = \frac{\omega_1}{\omega_1 + \omega_2}$. Note that

$$\varphi_1 + \varphi_2 = 2 \quad (\text{D.44})$$

From equations (D.41), (D.42), (D.43), and (D.44), we can solve for K_1 , K_2 , φ_1 , and φ_2 :

$$K_1 = \frac{2(1-\omega)(1-\theta_2)K^s}{(1-\omega)(1-\theta_2)(1+K^s) + \omega(1-\theta_1)(1-K^s)} \quad (\text{D.45})$$

$$K_2 = \frac{2\omega(1-\theta_1)K^s}{(1-\omega)(1-\theta_2)(1-K^s) + \omega(1-\theta_1)(1+K^s)}. \quad (\text{D.46})$$

It is clear that $\frac{\partial K_1}{\partial \omega} < 0$. In addition, $\frac{\partial \varphi_1}{\partial \omega} < 0$ due to equation (D.41), and $\frac{\partial \sigma_{e,1}^2}{\partial \omega} > 0$ because $\sigma_{e,1}^2 = (\frac{1}{K_1} - 1)\sigma_u^2$. Since κ_1 is increasing in K_1 , we have $\frac{\partial \kappa_1}{\partial \omega} < 0$. This completes the proof of Proposition 5. □

D.9 Interior Solutions and Corner Solutions

First, we list all of the conditions we need. We have proved that Kalman gains are as follows:

$$K_1 = \max\left\{\frac{2(1-\omega)(1-\theta_2)K_{ss}}{(1-\omega)(1-\theta_2)(1+K_{ss}) + \omega(1-\theta_1)(1-K_{ss})}, 0\right\} \quad (\text{D.47})$$

$$\equiv \max\{K_1^{pre}, 0\} \quad (\text{D.48})$$

$$K_2 = \max\left\{\frac{2\omega(1-\theta_1)K_{ss}}{(1-\omega)(1-\theta_2)(1-K_{ss}) + \omega(1-\theta_1)(1+K_{ss})}, 0\right\} \quad (\text{D.49})$$

$$\equiv \max\{K_2^{pre}, 0\} \quad (\text{D.50})$$

where we have used K_1^{pre} and K_2^{pre} to represent the first term in the two brackets to simplify the illustration. In addition, we have represented Kalman gains as function of policy

function φ_1 and φ_2 as in section C.3:

$$K_1 = \max \left\{ 1 - \frac{\lambda}{2\ln(2)(1-\theta_1)\varphi_1^2\sigma_u^2}, 0 \right\} \quad (\text{D.51})$$

$$K_2 = \max \left\{ 1 - \frac{\lambda}{2\ln(2)(1-\theta_2)\varphi_2^2\sigma_u^2}, 0 \right\}. \quad (\text{D.52})$$

And the monetary policy function follows

$$\frac{1}{2}(1-\theta_1)\omega\varphi_1K_1u - \frac{1}{2}(1-\theta_2)(1-\omega)\varphi_2K_2u = 0, \quad (\text{D.53})$$

Step 1: Suppose $\omega \neq 0, 1$. If attention problems for the sector 1 is at the corner solution, then the attention choice for sector 2 must be the corner solution as well: if $K_1 = 0$, then $K_2 = 0$. This is proven by contradiction: otherwise, it must be that $\varphi_2 = 0$, $K_2 \neq 0$ (fixing other parameters) as implied from the monetary policy (D.53). This contradicts with (D.52) because $\varphi_2 = 0$ implies $K_2 = 0$. Similarly, if $K_2 = 0$, K_1 must be 0.

Step 2: Suppose $0 < \omega < 1$, (D.47) and (D.49) imply that $K_1 > 0$ and $K_2 > 0$, leading to an interior solution.

Step 3: Suppose $\omega = 1$, then $K_1 = 0$ from (D.47), which implies $K_2 = 0$ from step 1. Similarly, when $\omega = 0$, $K_1 = K_2 = 0$.

Step 4: Suppose $\omega > 1$, then K_1^{pre} and K_2^{pre} must have opposite signs, either $K_1^{pre} < 0$ or $K_2^{pre} < 0$. In both cases, we must have $K_1 = K_2 = 0$ from step 1. Similarly, when $\omega < 1$, $K_1 = K_2 = 0$. Thus, the interior solution occurs when $0 < \omega < 1$, while a corner solution arises otherwise.

In the case of corner solution $K_1 = K_2 = 0$, the output gap becomes undetermined. From (C.31) and (C.32), any ϕ_x satisfies the two equations. For instance, suppose ϕ_x^1 is one equilibrium, then $\phi_x^1 + \zeta$ (for any white noise ζ) is another equilibrium. Thus, the output gap remains indeterminate.

D.10 Proof of Proposition 6

We follow the same procedure in D.3 to derive the optimal price index to replicate the allocations under Ramsey optimal policy. Then we show that $\omega_1^* < \omega_2^*$ if $\theta_1 < \theta_2$.

Step 1 : Given ϕ , all other endogenous variables are determined and can be derived. Specifically,

$$\begin{aligned} p + x &= \phi u \\ p_{1i}^* &= (\phi + 1)\mathbb{E}[u|s_{1i}] \\ p_{2i}^* &= (\phi - 1)\mathbb{E}[u|s_{2i}] \end{aligned}$$

The attention allocation is

$$K_k = \max \left\{ 1 - \left(\frac{2\lambda}{(1 - \theta_k)(\epsilon - 1)\varphi_k^2 \sigma_u^2} \right)^{\frac{1}{2}}, 0 \right\} = 1 - \left(\frac{1}{1 - \theta_k} \right)^{0.5} \frac{1 - K^s}{\varphi_k} \quad (\text{D.54})$$

where

$$\varphi_1 = \phi + 1, \quad \varphi_2 = 1 - \phi. \quad (\text{D.55})$$

The total welfare loss function can be decomposed into four components: (1) price dispersion within sectors due to sticky price within sectors $\mathbb{L}_{Calvo}$, (2), price dispersion due to dispersed within sector due to dispersed information $\mathbb{L}_{DS}$ (3) output gap $\mathbb{L}_x$ and (4) relative price gap across sectors $\mathbb{L}_{rp}$. The total welfare loss function is

$$\mathbb{L}_{total} = \mathbb{L}_{Calvo} + \mathbb{L}_{DS} + \mathbb{L}_x + \mathbb{L}_{rp}$$

where

$$\begin{aligned} \mathbb{L}_{Calvo} &= \frac{\epsilon}{2} \left[(1 - \theta_1) \theta_1 (\phi + 1)^2 K_1^2 \sigma_u^2 + (1 - \theta_2) \theta_2 (\phi - 1)^2 K_2^2 \sigma_u^2 \right] \\ \mathbb{L}_{DS} &= \frac{\epsilon}{2} \left[(1 - \theta_1) (\phi + 1)^2 K_1^2 \sigma_{\epsilon 1}^2 + (1 - \theta_2) (\phi - 1)^2 K_2^2 \sigma_{\epsilon 2}^2 \right] \\ \mathbb{L}_x &= \left[\left(1 - \frac{1 - \theta_1}{2} K_1 - \frac{1 - \theta_2}{2} K_2 \right) \phi + \frac{1 - \theta_2}{2} K_2 - \frac{1 - \theta_1}{2} K_1 \right]^2 \sigma_u^2 \\ \mathbb{L}_{rp} &= \frac{\eta}{4} \sigma_u^2 \{ (1 - \theta_1) K_1 (\phi + 1) - (1 - \theta_2) K_2 (\phi - 1) - 2 \}^2 \end{aligned}$$

These expressions can be simplified as follows:

$$\begin{aligned}\mathbb{L}_{Calvo} &= \frac{\epsilon}{2} \sigma_u^2 \left[(1 - \theta_1) \theta_1 [\phi + Q_1]^2 + (1 - \theta_2) \theta_2 [\phi + Q_2]^2 \right] \\ \mathbb{L}_{DS} &= \frac{\epsilon}{2} \sigma_u^2 \left[(1 - \theta_1) [\phi + Q_1] \left(\frac{1}{1 - \theta_1} \right)^{\frac{1}{2}} (1 - K^s) + (1 - \theta_2) [\phi + Q_2] \left(\frac{1}{1 - \theta_2} \right)^{\frac{1}{2}} (K^s - 1) \right] \\ \mathbb{L}_x &= \frac{\sigma_u^2}{4} [(\theta_1 + \theta_2) \phi - (1 - \theta_1) Q_1 - (1 - \theta_2) Q_2]^2 \\ \mathbb{L}_{rp} &= \frac{\eta}{4} \sigma_u^2 \{ (1 - \theta_1) (\phi + Q_1) - (1 - \theta_2) (\phi + Q_2) - 2 \}^2\end{aligned}$$

where $Q_1 = 1 - \left(\frac{1}{1 - \theta_1} \right)^{\frac{1}{2}} (1 - K^s)$, and $Q_2 = -1 + \left(\frac{1}{1 - \theta_2} \right)^{\frac{1}{2}} (1 - K^s)$.

Step 2 Second, we determine the optimal ϕ that minimizes the total welfare loss. To do this, we calculate the first-order derivative of the total welfare loss, $\mathbb{L}_{total}$, with respect to ϕ and set it equal to zero:

$$\frac{\partial \mathbb{L}_{total}}{\partial \phi} = 0$$

The optimal ϕ^* that minimize welfare loss is given by:

$$\phi^* = \frac{D_2}{D_1}$$

where D_1 and D_2 are defined as follows:

$$\begin{aligned}D_1 &= \frac{\epsilon}{2} \{ 2(1 - \theta_1) \theta_1 + 2(1 - \theta_2) \theta_2 \} \\ &\quad + \frac{1}{2} (\theta_1 + \theta_2)^2 + \frac{\eta}{2} (\theta_2 - \theta_1)^2 \\ D_2 &= -\frac{\epsilon}{2} \left\{ 2(1 - \theta_1) \theta_1 Q_1 + 2(1 - \theta_2) \theta_2 Q_2 + (1 - \theta_1)^{\frac{1}{2}} (1 - K^s) + (1 - \theta_2)^{\frac{1}{2}} (K^s - 1) \right\} \\ &\quad + \frac{1}{2} (\theta_1 + \theta_2) [(1 - \theta_1) Q_1 + (1 - \theta_2) Q_2] - \frac{\eta}{2} (\theta_2 - \theta_1) [(1 - \theta_1) Q_1 - (1 - \theta_2) Q_2 - 2]\end{aligned}$$

Step 3 Third, we find the optimal price index to implement the optimal monetary policy. Suppose the optimal price index is:

$$\omega_1 p_1 + \omega_2 p_2 = 0$$

Then,

$$\frac{\omega_1}{\omega_2} = \frac{\theta_2 - 1}{1 - \theta_1} \frac{D_2 + Q_2 D_1}{D_2 + Q_1 D_1}$$

The price weight ω_1 and ω_2 are not unique. We focus on the parameter region ensuring that $\omega_1/\omega_2 > 0$. Without the loss of generality, we focus on positive ω_1 and ω_2 , and choose

$$\begin{aligned}\omega_1 &= (\theta_2 - 1) \left\{ \frac{\epsilon}{2} \left[2(1 - \theta_1) \theta_1 (Q_2 - Q_1) + (K_s - 1) \left[(1 - \theta_1)^{\frac{1}{2}} - (1 - \theta_2)^{\frac{1}{2}} \right] \right] \right. \\ &\quad \left. + \frac{1}{2} (\theta_1 + \theta_2) [(\theta_1 + 1) Q_2 + (1 - \theta_1) Q_1] + \frac{\eta}{2} (\theta_2 - \theta_1) [(1 - \theta_1) (Q_2 - Q_1) + 2] \right\} \\ \omega_2 &= (1 - \theta_1) \left\{ \frac{\epsilon}{2} \left[2(1 - \theta_2) \theta_2 (Q_1 - Q_2) + (K_s - 1) \left[(1 - \theta_1)^{\frac{1}{2}} - (1 - \theta_2)^{\frac{1}{2}} \right] \right] \right. \\ &\quad \left. + \frac{1}{2} (\theta_1 + \theta_2) [(\theta_2 + 1) Q_1 + (1 - \theta_2) Q_2] + \frac{\eta}{2} (\theta_2 - \theta_1) [(1 - \theta_2) (Q_2 - Q_1) + 2] \right\}\end{aligned}$$

Step 4: comparing the optimal weight The difference of $\omega_1 - \omega_2$ is given by:

$$\begin{aligned}\omega_2 - \omega_1 &= \frac{\epsilon}{2} \left\{ 2(1 - \theta_1) (1 - \theta_2) (\theta_2 - \theta_1) (Q_1 - Q_2) + (1 - K^s) \left[(1 - \theta_2)^{\frac{1}{2}} - (1 - \theta_1)^{\frac{1}{2}} \right] (2 - \theta_1 - \theta_2) \right\} \\ &\quad + \frac{1}{2} (\theta_1 + \theta_2) [2(1 - \theta_1) Q_1 + 2(1 - \theta_2) Q_2] \\ &\quad + \frac{\eta}{2} (\theta_2 - \theta_1) [2(1 - \theta_1) (1 - \theta_2) (Q_2 - Q_1) + 4 - 2\theta_1 - 2\theta_2]\end{aligned}$$

where

$$Q_1 = 1 - \left(\frac{1}{1 - \theta_1} \right)^{\frac{1}{2}} (1 - K^s), Q_2 = -1 + \left(\frac{1}{1 - \theta_2} \right)^{\frac{1}{2}} (1 - K^s)$$

For simplicity, denote $q_1 \equiv 1 - \theta_1 \in (0, 1)$, $q_2 \equiv 1 - \theta_2 \in (0, 1)$. Given $\theta_1 < \theta_2$, we have $q_1 > q_2$. Also, denote $K \equiv 1 - K^s \in (0, 1)$. Thus,

$$\begin{aligned}Q_1 &= 1 - q_1^{-\frac{1}{2}} K \\ Q_2 &= -1 + q_2^{-\frac{1}{2}} K \\ Q_1 - Q_2 &= 2 - (q_1^{-\frac{1}{2}} + q_2^{-\frac{1}{2}}) K\end{aligned}$$

and

$$\begin{aligned}\omega_2 - \omega_1 = & \frac{\epsilon}{2} \left\{ 2q_1q_2(q_1 - q_2) \left[2 - (q_1^{-\frac{1}{2}} + q_2^{-\frac{1}{2}})K \right] + K(q_2^{\frac{1}{2}} - q_1^{\frac{1}{2}})(q_1 + q_2) \right\} \\ & + \frac{1}{2}(2 - q_1 - q_2) \left[2q_1(1 - q_1^{-\frac{1}{2}}K) + 2q_2(-1 + q_2^{-\frac{1}{2}}K) \right] \\ & + \frac{\eta}{2}(q_1 - q_2) \left\{ 2q_1q_2 \left[-2 + (q_1^{-\frac{1}{2}} + q_2^{-\frac{1}{2}})K \right] + 2(q_1 + q_2) \right\}\end{aligned}$$

These expressions can be simplified as follows:

$$\omega_2 - \omega_1 = -X(q_1^{\frac{1}{2}} - q_2^{\frac{1}{2}})K + Y$$

where

$$X = \left[\epsilon q_1^{\frac{1}{2}} q_2^{\frac{1}{2}} (q_1^{\frac{1}{2}} + q_2^{\frac{1}{2}})^2 + \frac{\epsilon}{2}(q_1 + q_2) \right] + (2 - q_1 - q_2) - \eta q_1^{\frac{1}{2}} q_2^{\frac{1}{2}} (q_1^{\frac{1}{2}} + q_2^{\frac{1}{2}})^2$$

$$Y = (q_1 - q_2) \left[2\epsilon q_1q_2 + (2 - q_1 - q_2) + \eta(q_1 + q_2 - 2q_1q_2) \right]$$

Note that if $\epsilon > \eta$, then $X > 0$. Therefore $\omega_1 - \omega_2 > 0$ holds if and only if

$$(q_1^{\frac{1}{2}} - q_2^{\frac{1}{2}})K > \frac{Y}{X} = \frac{(q_1 - q_2) \left[2\epsilon q_1q_2 + (2 - q_1 - q_2) + \eta(q_1 + q_2 - 2q_1q_2) \right]}{(\epsilon - \eta)q_1^{\frac{1}{2}}q_2^{\frac{1}{2}}(q_1^{\frac{1}{2}} + q_2^{\frac{1}{2}})^2 + \frac{\epsilon}{2}(q_1 + q_2) + (2 - q_1 - q_2)}$$

That is

$$K > K^* \equiv \frac{(q_1^{\frac{1}{2}} + q_2^{\frac{1}{2}}) \left[2\epsilon q_1q_2 + (2 - q_1 - q_2) + \eta(q_1 + q_2 - 2q_1q_2) \right]}{(\epsilon - \eta)q_1^{\frac{1}{2}}q_2^{\frac{1}{2}}(q_1^{\frac{1}{2}} + q_2^{\frac{1}{2}})^2 + \frac{\epsilon}{2}(q_1 + q_2) + (2 - q_1 - q_2)}$$

Recall that $K^s \equiv 1 - K$. Therefore, the previous condition is equivalent to:

$$K^s < 1 - \frac{(q_1^{\frac{1}{2}} + q_2^{\frac{1}{2}}) \left[2\epsilon q_1q_2 + (2 - q_1 - q_2) + \eta(q_1 + q_2 - 2q_1q_2) \right]}{(\epsilon - \eta)q_1^{\frac{1}{2}}q_2^{\frac{1}{2}}(q_1^{\frac{1}{2}} + q_2^{\frac{1}{2}})^2 + \frac{\epsilon}{2}(q_1 + q_2) + (2 - q_1 - q_2)}$$

In sum, $\omega_1 > \omega_2$ given $\theta_1 < \theta_2$ if and only if $K^s < \bar{K}^{alt}$, where

$$\bar{K}^{alt} = 1 - \frac{(q_1^{\frac{1}{2}} + q_2^{\frac{1}{2}}) [2\epsilon q_1 q_2 + (2 - q_1 - q_2) + \eta(q_1 + q_2 - 2q_1 q_2)]}{(\epsilon - \eta) q_1^{\frac{1}{2}} q_2^{\frac{1}{2}} (q_1^{\frac{1}{2}} + q_2^{\frac{1}{2}})^2 + \frac{\epsilon}{2}(q_1 + q_2) + (2 - q_1 - q_2)} \quad (D.56)$$

or equivalently

$$\bar{K}^{alt} = 1 - \frac{[(1 - \theta_1)^{\frac{1}{2}} + (1 - \theta_2)^{\frac{1}{2}}] [2\epsilon(1 - \theta_1)(1 - \theta_2) + (\theta_1 + \theta_2) + \eta(2 - \theta_1 - \theta_2 - 2(1 - \theta_1)(1 - \theta_2))]}{(\epsilon - \eta)(1 - \theta_1)^{\frac{1}{2}}(1 - \theta_2)^{\frac{1}{2}} [(1 - \theta_1)^{\frac{1}{2}} + (1 - \theta_2)^{\frac{1}{2}}]^2 + \frac{\epsilon}{2}(2 - \theta_1 - \theta_2) + (\theta_1 + \theta_2)}$$

E Dynamic Model

In this section, we provide a dynamic multi-sector model under rational inattention.

Households We assume households have complete information. The representative household solves a lifetime utility maximization problem:

$$\max_{C_t, L_{k,t}, B_{t+1}} E_0 \sum_{t=0}^{\infty} \beta^t \left(\log(C_t) - \sum_{k=1}^2 L_{k,t} \right), \quad (E.1)$$

$$\text{s.t. } P_t C_t + Q_t B_{t+1} = B_t + \sum_{k=1}^2 W_{k,t} L_{k,t} + \sum_{k=1}^2 \Pi_{k,t} + T_t, \quad (E.2)$$

where P_t is the price level of composite good; Q_t is the nominal one-period risk-free bond price; $W_{k,t}$ is the sectoral nominal wage; $\Pi_{k,t}$ is sectoral profit; and T_t represents lump-sum taxes/transfers. Solving the household problem gives rise to the first-order conditions:

$$\frac{1}{C_t} \frac{W_{k,t}}{P_t} = 1 \quad (E.3)$$

$$Q_t = \beta E_t \left(\frac{C_t P_t}{C_{t+1} P_{t+1}} \right) \quad (E.4)$$

The aggregate consumption C_t is a CES aggregator of sectoral consumption $C_{k,t}$ and the sectoral consumption $C_{k,t}$ is a CES aggregate of individual consumption $C_{ki,t}$. Formally,

$$C_t = \left[\sum_{k=1}^2 n_k^{1/\eta} C_{k,t}^{(\eta-1)/\eta} \right]^{\eta/(\eta-1)}, \quad C_{k,t} = \left[n_k^{-1/\epsilon} \int_0^{n_k} C_{ki,t}^{(\epsilon-1)/\epsilon} di \right]^{\epsilon/(\epsilon-1)}$$

Optimizing the allocation of the consumption within and across the sectors, we yield the demand

for variety i within sector k ($C_{ki,t}$) and across sectors ($C_{k,t}$):

$$C_{ki,t} = \left(\frac{P_{ki,t}}{P_{k,t}} \right)^{-\epsilon} \frac{1}{n_k} C_{k,t}, \quad C_{k,t} = \left(\frac{P_{k,t}}{P_t} \right)^{-\eta} n_k C_t. \quad (\text{E.5})$$

where the implied aggregate and sectoral prices are

$$P_t \equiv \left(\sum_{k=1}^2 n_k P_{k,t}^{1-\eta} \right)^{1/(1-\eta)}, \quad P_{k,t} \equiv \left(\int_0^1 P_{ki,t}^{1-\epsilon} di \right)^{1/(1-\epsilon)} \quad (\text{E.6})$$

Firms We assume a continuum of monopolistically competitive firms in each sector k . Each firm i produces output using the production function $Y_{ki,t} = e^{a_{k,t}} L_{ki,t}$, where $a_{k,t}$ is the log productivity in sector k . We assume $a_{k,t}$ evolves according to an autoregressive process: $a_{k,t} = \rho a_{k,t-1} + \epsilon_{k,t}$, where $\epsilon_{k,t}$ is an i.i.d. shock in sector k .

First, conditional on observed signals, we solve for the optimal price-setting. Similar to the static model, by taking a second-order approximation of the firms' profit function, we have

$$V(p_{ki,t}^*) = E_{ki,t} \sum_{j=0}^{\infty} (\beta \theta_k)^j \left(-\frac{\epsilon-1}{2} p_{ki,t}^{*2} + (\epsilon-1) p_{ki,t}^* \varphi_{k,t+j} \right) + t.i.p.s. \quad (\text{E.7})$$

where *t.i.p.s* includes the terms that have no impact on the optimal price and

$$\varphi_{k,t+j} \equiv p_{t+j} + y_{t+j} - a_{k,t+j} \quad (\text{E.8})$$

is the marginal cost. The first-order condition for the optimal price $p_{ki,t}^*$ is

$$p_{ki,t}^* = E_{ki,t}(p_{ki,t}^\diamond | I_{ki,t}), \quad (\text{E.9})$$

and $p_{ki,t}^\diamond$ is the profit-maximizing price of firm i in sector k in period t :

$$p_{ki,t}^\diamond = (1 - \beta \theta) \sum_{j=0}^{\infty} (\beta \theta)^j E_t^{FI} \varphi_{k,t+j} = (1 - \beta \theta) \sum_{k=0}^{\infty} (\beta \theta)^j E_t^{FI} (p_{t+j} + y_{t+j} - a_{k,t+j}) \quad (\text{E.10})$$

Where E_t^{FI} is the expectation operator under full information. Rewrite equation (E.10) in recursive form:

$$p_{1i,t}^\diamond = (1 - \beta \theta_1) \left(\sum_{k=1}^2 n_k p_{k,t} \right) + (1 - \beta \theta_1)(x_t + \zeta_t) + \beta \theta_1 E_t^{FI}(p_{1i,t+1}^\diamond), \quad (\text{E.11})$$

$$p_{2i,t}^\diamond = (1 - \beta \theta_2) \left(\sum_{k=1}^2 n_k p_{k,t} \right) + (1 - \beta \theta_2)(x_t - \zeta_t) + \beta \theta_2 E_t^{FI}(p_{2i,t+1}^\diamond), \quad (\text{E.12})$$

where $\xi_t \equiv \frac{1}{2}(a_{2t} - a_{1t})$ is the relative productivity as in the static model, and $x_t = y_t - y_t^N$ is output gap with y^N representing the natural level of output.²¹

In stage 2, conditional on the optimal prices, firms choose the optimal amount of information to maximize their profit subject to the information flow constraint. Following [Mackowiak and Wiederholt \(2009\)](#), we transform the profit maximization problem to the profit loss minimization problem. We define the profit loss from the information friction as follows:

$$V(p_{ki,t}^\diamond) - V(p_{ki,t}^*) = E_{ki,t} \sum_{j=0}^{\infty} (\beta\theta_k)^j \frac{\epsilon - 1}{2} (p_{ki,t}^{*2} - p_{ki,t}^{\diamond 2} - 2\varphi_{k,t+j}(p_{ki,t}^* - p_{ki,t}^\diamond)) \quad (\text{E.13})$$

$$= \frac{\epsilon - 1}{2} E_{ki,t} \sum_{j=0}^{\infty} (\beta\theta_k)^j (p_{ki,t}^* - p_{ki,t}^\diamond)(p_{ki,t}^* + p_{ki,t}^\diamond - 2\varphi_{k,t+j}) \quad (\text{E.14})$$

$$= \frac{\epsilon - 1}{2} E_{ki,t} \sum_{j=0}^{\infty} (\beta\theta_k)^j (p_{ki,t}^* - p_{ki,t}^\diamond)^2 \quad (\text{E.15})$$

$$= \frac{\epsilon - 1}{2} \frac{1}{1 - \beta\theta_k} E_{ki,t} (p_{ki,t}^* - p_{ki,t}^\diamond)^2 \quad (\text{E.16})$$

where from the second equality to the third equality we have used [\(E.10\)](#).

Similar to the static setting, we have defined the relative productivity $\xi_t \equiv \frac{1}{2}(a_{2,t} - a_{1,t}) = \rho\xi_{t-1} + u_t$, where $u_t \equiv \frac{1}{2}(\epsilon_{2,t} - \epsilon_{1,t})$. In the dynamic setting, the unknown states include the whole history of shocks. For a length of truncation L , the state space is $u^t \equiv (u_t, u_{t-1}, \dots, u_{t-(L-1)})'$. The firm wants to keep track of the ideal price $p_{ki,t}^\diamond$, which is a Gaussian process. By Wold's theorem, we can decompose it into its $MA(\infty)$ representation: $p_{ki,t}^\diamond = \Phi_k(L)u_t$, in which $\Phi_k(\cdot)$ are lag polynomials. The law of motion of u^t can be written as

$$u^t = Au^{t-1} + Qu_t, \quad (\text{E.17})$$

which is determined by a $L \times 1$ vector $Q = (1 \ 0 \ 0 \ \dots \ 0)'$ and a $L \times L$ matrix

$$A = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \dots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

Firms do not know the true state, but observe a noisy signal of the following structure, proved by

²¹ $y_t^N = \frac{a_{1,t} + a_{2,t}}{2}$ in equilibrium.

Afrouzi and Yang (2021)

$$s_{ki,t} = \mathbf{y}'_{ki} u^t + e_{ki,t} \quad (\text{E.18})$$

with $e_{ki,t} \sim N(0, \sigma_{e,ki,t}^2)$ and $\mathbf{y}_{ki}$ is the loading of signal on state space u^t . Note that although $\sigma_{e,ki,t}^2$ has subscript t to represent its potential variation over time, it will be a constant in the stationary equilibrium considered in this analysis.

Now, we can formalize the attention problem as the following linear-quadratic-Gaussian form:

$$\min_{\Sigma_{ki,t|t} \in \mathcal{S}_+^L} \frac{\epsilon - 1}{2} \frac{1}{1 - \beta \theta_k} E_{ki,t} (p_{ki,t}^* - p_{ki,t}^\diamond)^2 + \frac{\lambda}{2} \log_2 \left(\frac{|\Sigma_{ki,t|t-1}|}{|\Sigma_{ki,t|t}|} \right), \quad (\text{E.19})$$

subject to the optimal price setting rule (4.3), and the constraints on the prior and the posterior uncertainty about u^t :

$$\Sigma_{ki,t+1|t} = A \Sigma_{t|t} A' + Q Q' \quad \forall t \geq 0, \quad (\text{E.20})$$

$$\Sigma_{ki,t|t-1} - \Sigma_{t|t} \succeq 0, \quad \forall t \geq 0, \quad (\text{E.21})$$

$$0 \prec \Sigma_{ki,0|-1} \prec \infty \quad \text{given}, \quad (\text{E.22})$$

where $|\cdot|$ is the determinant operator, $\succeq$ denotes positive semidefiniteness, $\Sigma_{ki,t|t} \equiv \text{var}(u^t | s_{ki}^t)$, $\Sigma_{ki,t|t-1} \equiv \text{var}(u^t | s_{ki}^{t-1})$, $s_{ki}^t = \{s_{ki,1}, s_{ki,2}, \dots, s_{ki,t}\}$ denote the sequence of all signals that firm i has received up to period t , λ is the marginal information processing cost, $\Sigma_{ki,t|t} \equiv \text{var}(u^t | s_{ki}^t)$ is a L -dimensional symmetric positive semidefinite cone, and $\Sigma_{ki,0|-1}$ is the prior variance-covariance matrix. Following Afrouzi et al. (2024), we have assumed a linear attention cost function, $f(\kappa_{ki,t}) = \kappa_{ki,t}$, in this dynamic setting. This assumption simplifies the solution of the model, as it implies that prior uncertainty $\Sigma_{ki,t|t-1}$ does not affect the choice of the optimal $\Sigma_{ki,t|t}$ when firms come to reset prices. And therefore the objective function for the information acquisition problem can be simplified to (E.19).

Other Equilibrium Conditions The log level price in sector k evolves as the following

$$p_{k,t} = (1 - \theta_k) p_{k,t}^* + \theta_k p_{k,t-1} \quad (\text{E.23})$$

where $p_{1,t}^*$ and $p_{2,t}^*$ are the average optimal prices that re-optimizing firm i in sectors 1 and 2 would set:

$$p_{k,t}^* = \int_0^1 p_{ki,t}^* di, \quad (\text{E.24})$$

Given the definition of inflation, we have

$$\pi_{k,t} = (1 - \theta_k)(p_{k,t}^* - p_{k,t-1}) \quad (\text{E.25})$$

In addition, the relative output in sector k is defined as $y_{R,kt} = y_{k,t} - y_t$. Then, the law of motion of $y_{R,kt}$ is

$$y_{R,kt} = -\eta \left(\pi_{k,t} - \sum_{k=1}^2 n_k \pi_{k,t} \right) + y_{R,kt-1}. \quad (\text{E.26})$$

The relative output gap is

$$\tilde{y}_{R,kt} = y_{R,kt} - \eta \left(a_{k,t} - \sum_{k=1}^{N_k} n_{N_k} a_{k,t} \right). \quad (\text{E.27})$$

The monetary policy follows an inflation stabilization policy:

$$\omega \pi_{1,t} + (1 - \omega) \pi_{2,t} = 0. \quad (\text{E.28})$$

The welfare loss function is:

$$E\mathbb{L} = E \left[\sum_{k=1}^2 (\epsilon n_k \text{var}_i(p_{ki,t})) + x_t^2 + \eta^{-1} \sum_{k=1}^2 n_k \tilde{y}_{R,kt}^2 \right]. \quad (\text{E.29})$$

E.1 Equilibrium under Optimal Inflation Stabilization Policy

The equilibrium under the optimal inflation stabilization policy is defined as the central bank optimally choosing ω such that the endogenous variables $p_{ki,t}^*$, $p_{ki,t}^\diamond$, $p_{k,t}$, p_t , $\pi_{k,t}$, π_t , $y_{R,kt}$, $\tilde{y}_{R,kt}$, $\Sigma_{ki,t|t}$, and x_t , given the information structure (E.18), satisfy the conditions in (E.9), (E.11), (E.12), (E.19), (E.24), (E.25), (E.26), (E.27), (E.28), and the welfare loss (E.29) is minimized.

E.2 Solving the Two-sector NK Model under Rational Inattention

Following Afrouzi and Yang (2021), the attention problem can be characterized by the following conditions:

$$\hat{\lambda} \Sigma_{ki,t|t}^{-1} - \Lambda_t = \Omega + A'(\hat{\lambda} \Sigma_{ki,t+1|t}^{-1} - \Lambda_{ki,t+1})A, \quad \forall t \geq 0, \quad (\text{E.30})$$

$$\Lambda_t(\Sigma_{ki,t|t-1} - \Sigma_{ki,t|t}) = 0, \Lambda_t \succeq 0, \Sigma_{ki,t|t-1} - \Sigma_{ki,t|t} \succeq 0, \quad \forall t \geq 0, \quad (\text{E.31})$$

$$\Sigma_{ki,t+1|t} = A \Sigma_{ki,t|t} A' + Q Q', \quad (\text{E.32})$$

where Λ_t are the Lagrange multipliers on the no-forgetting constraints (E.21), $\hat{\lambda} \equiv \frac{\lambda}{\ln(2)} / (\frac{\epsilon-1}{2} \frac{1}{1-\beta\theta_k})$ is the normalized information-processing cost, $\Omega \equiv H_k H_k'$, and H_k represent the policy function for $p_{ki,t}^\diamond$. Notice that given the policy function H_k , we can solve for the optimal posterior variance-covariance matrix $\Sigma_{ki,t|t}$. Once again, we focus on a stationary information structure when $\Sigma_{ki,t|t}$ converges. And $\Sigma_{ki,t|t} = \Sigma_{kj,t|t} = \Sigma_{k,t|t}$, $\Sigma_{ki,t|t-1} = \Sigma_{kj,t|t-1} = \Sigma_{k,t|t-1}$ in equilibrium because of the symmetric feature within a sector.

Before showing how to obtain the solutions for the policy function and attention problem through iteration, we must adjust the state space to accommodate the unit root in prices. We follow Afrouzi and Yang (2021) and implement the following transformation. First, define $\tilde{u}_t$ such that $(1 - \mathbf{L})\tilde{u}_t = u_t$ with $\mathbf{L}$ as the lag operator. In other words, $\tilde{u}_t = (1 - \mathbf{L})^{-1}u_t$ and $\tilde{u}_t$ becomes a unit root process. Then define $\tilde{u}^t \equiv (\tilde{u}_t, \tilde{u}_{t-1}, \dots, \tilde{u}_{t-(L-1)})'$. It is also easy to see that

$$\tilde{u}^t - M' \tilde{u}^t = u^t.$$

Putting it more compactly, it is

$$u^t = (I - M') \tilde{u}^t \quad \text{or} \quad \tilde{u}^t = (I - M')^{-1} u^t,$$

where I is the $L \times L$ identity matrix and M is a $L \times L$ shift matrix

$$M = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \dots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

Thus, $\tilde{u}^t$ follows

$$\tilde{u}^t = \tilde{A} \tilde{u}^{t-1} + Q u_t, \tag{E.33}$$

where Q remains and

$$\tilde{A} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \dots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

Our analysis above also applies to the new state space $\tilde{u}^t$. To simplify the notation, we will not redefine each related term associated with $\tilde{u}^t$. With slight abuse of notation, all terms below, such as $K_1, K_2, \mathbf{y}_1, \mathbf{y}_2, \Sigma_{1i,t|t-1}, \Sigma_{1i,t|t}, \Sigma_{2i,t|t-1}, \Sigma_{2i,t|t}$ are taken with respect to the new state space $\tilde{u}^t$.

Now, we guess $p_{1i,t}^\diamond = H_1' \tilde{u}^t$, $p_{2i,t}^\diamond = H_2' \tilde{u}^t$ and $x_t = H_y' u^t$ and iterate the policy functions until they converge. The iteration procedure is stated as follows.

- With our guess, we first obtain $p_{1,t}^*$ as:

$$\begin{aligned} p_{1,t}^* &= \int p_{1,t}^*(i) di = \int E_{1i,t}[p_{1i,t}^\diamond] di, \\ &= H_{1,(n-1)}' X_{1,(n)} \tilde{u}^t, \end{aligned}$$

where $X_{1,(n)} = \sum_{j=0}^{\infty} \left[(I - K_{1,(n)} \mathbf{y}_{1,(n)}') \tilde{A} \right]^j K_{1,(n)} \mathbf{y}_{1,(n)}' M'^j$, and $\mathbf{y}_{1,(n)}$ is the loading of signal on u^t for firms in sector 1, and $K_{1,(n)}$ is the Kalman gain of firms in sector 1, which is defined as $K_{1,(n)} \equiv \Sigma_{1i,t|t-1} \mathbf{y}_{1,(n)}' (\mathbf{y}_{1,(n)}' \Sigma_{1i,t|t-1} \mathbf{y}_{1,(n)} + \sigma_{e,1,t}^2)^{-1}$ and $\sigma_{e,1,t}^2$ is the variance of the noise in signal for firms in sector 1. Note that $K_{1,(n)}$ and $\mathbf{y}_{1,(n)}$ are obtained given the pricing policy function $H_{1,(n-1)}$, which makes use of the algorithm [Afrouzi and Yang \(2021\)](#) provides. Note that the subscript (n) refer to the $(n)^{th}$ iteration.

- Then we obtain $p_{1,t} = H_{p,1}' \tilde{u}^t$ and $\pi_{1,t} = H_{\pi,1}' u^t$:

$$\begin{aligned} p_{1,t} &= \theta_1 p_{1,t-1} + (1 - \theta_1) p_{1,t}^* \\ &= (1 - \theta_1) \sum_{j=0}^{\infty} \theta_1^j p_{1,t-j}^* \\ &= (1 - \theta_1) H_{1,(n-1)}' X_{1,(n)} \sum_{j=0}^{\infty} (\theta_1 M')^j \tilde{u}^t \\ &\equiv H_{p,1}' \tilde{u}^t \end{aligned}$$

and thus,

$$\begin{aligned} \pi_{1,t} &= p_{1,t} - p_{1,t-1} \\ &= H_{p,1}' (I - M')^{-1} (I - M') u^t \equiv H_{\pi,1}' u^t \end{aligned}$$

This implies that $H_{p,1} = H_{\pi,1}$.

- We use the monetary policy to obtain $p_{2,t} = H_{p,2}' \tilde{u}^t$, and $\pi_{2,t} = H_{\pi,2}' u^t$:

$$H_{p,2} = H_{\pi,2} = -\frac{\omega}{1 - \omega} H_{\pi,1}$$

- We use pricing equation for sector 1 ([E.11](#)) to update H_1 :

$$H_{1,(n)} = (1 - \beta\theta_1)(0.5H_{p,1} + 0.5H_{p,2} + (I - M)(H_{y,(n-1)} + H_u)) + \beta\theta_1(I - M)^{-1} \tilde{A}' H_{1,(n-1)}$$

where we have used $u_t = H'_u u^t$.

- We use pricing equation for sector 2 (E.12) to update H_2 :

$$H_{2,(n)} = (1 - \beta\theta_2)(0.5H_{p,1} + 0.5H_{p,2} + (I - M)(H_{y,(n-1)} - H_u)) + \beta\theta_2(I - M)^{-1}\tilde{A}'H_{2,(n-1)}.$$

- We then update H_y :

$$H_{y,(n)} = H_{y,(n-1)} + H_{p,2} - (1 - \theta_2) \sum_{j=0}^{\infty} (\theta_2 M)^j X'_{2,(n)} H_{2,(n-1)},$$

where $X_{2,(n)} = \sum_{j=0}^{\infty} \left[(I - K_{2,(n)} \mathbf{y}'_{2,(n)}) \tilde{A} \right]^j K_{2,(n)} \mathbf{y}'_{2,(n)} M'^j$, and $\mathbf{y}_{2,(n)}$ is the loading of signal on u^t for firms in sector 2. They are obtained in the same way as $X_{1,(n)}$ and $\mathbf{y}_{1,(n)}$.

- With newly updated $H_{1,(n)}$, $H_{2,(n)}$ and $H_{y,(n)}$, we update the the Kalman gains and signal loadings $K_{1,(n+1)}$, $K_{2,(n+1)}$, $\mathbf{y}_{1,(n+1)}$ and $\mathbf{y}_{2,(n+1)}$ following Afrouzi and Yang (2021), allowing us to proceed to the next iteration. The iteration continues until H_1 , H_2 and H_y converge.

Note that to obtain the output gap we use the only remaining equation $p_{2,t} = (1 - \theta_2) \sum_{j=0}^{\infty} (\theta_2 M)^j X'_2 H_2 \tilde{u}^t$ which is derived similar to $p_{1,t}$. We add the output gap to both sides to ensure convergence of the output gap.²² Once this equation for $p_{2,t}$ holds, the output gap also converges. After the iteration converges, we obtain the policy function for relative output H_R as $H_R = -\frac{\eta}{2}(H_{\pi 1} - H_{\pi 2}) + MH_R$. The policy function for the relative output gap is then derived as $H_R + \eta H_u$, given that $\tilde{y}_{R,t} = y_{R,t} - \frac{\eta}{2}(a_{1,t} - a_{2,t})$.

²²Marcet and Singleton (1999) applies a similar logic to solve the portfolio choice problem.

F Appendix for Data Analysis

F.1 Data Source

10-K filings: The body of the 10-K filings is retrieved from The Notre Dame Software Repository for Accounting and Finance (SRAF). The SRAF is a repository of financial and accounting data maintained by the University of Notre Dame, and it is available at: <https://sraf.nd.edu/data/stage-one-10-x-parse-data/>

Firm-level data: The firm-level data is retrieved from Compustat. We use data on sales to weight each firm when we aggregate into sectors.

F.2 Dataset Construction

We adopt the text-based approach from [Song and Stern \(2022\)](#) to gather attention data by generating word counts on various topics. Our data is derived from the body of 10-K filings from the Securities and Exchange Commission (SEC) spanning the years 1994 to 2019. These filings provide a comprehensive overview of a company’s financial performance, including information on its history, organizational structure, audited financial statements, and risk factors, among other details. All companies that trade their securities on a US exchange and are subject to Section 13 or 15(d) of the US Securities Exchange Act of 1934 are required to file these reports. The text versions of these filings are accessed via the Software Repository for Accounting and Finance at the University of Notre Dame, as originally proposed by [Loughran and McDonald \(2016\)](#).

To calculate word counts, we employ the method proposed by [Loughran and McDonald \(2011\)](#), also used in [Song and Stern \(2022\)](#). This process involves parsing the filings by removing numbers and “stop words” (e.g., connectors) and mapping each word to a dictionary containing all words in the sample. We use the stop words list provided by [Loughran and McDonald \(2011\)](#)²³, which may differ from the one used in [Song and Stern \(2022\)](#), as they do not provide a description of their stop words list. Subsequently, we apply the same topics and keywords used by [Song and Stern \(2022\)](#) to construct measures of attention intensity, which represents the number of times the topic is mentioned per thousand words in the filing.

We work with an unbalanced panel dataset containing 212,000 observations for 35,000 firms. However, we only have data for 753 sectors (no time series), which correspond to those with data on the FPA. Therefore, we aggregate the firm-level panel data into sector-level data. To do so, we match the firm data with Compustat to obtain individual sales data. We then aggregate the firm-level data into sector-level data by calculating simple means of observations within each

²³<https://sraf.nd.edu/textual-analysis/stopwords/>

firm. Subsequently, we compute the sales-weighted average of firms within each sector. The final dataset comprises a panel of sector-topic observations, where each sector has a unique FPA value.

Two factors contribute to the loss of approximately one-third of the observations during this process. First, there is an imperfect match between CIK (SEC's firm identifiers) and GVKEY (Compustat's firm identifiers). Second, some firms do not report their 6-digit NAICS in Compustat, as some entries either have missing values or only report 2-digit or 4-digit NAICS. As a result, the dataset is reduced to 156,000 observations, which further decreases to 136,000 after selecting the dates for our analysis (1994-2019). Additional data loss occurs after merging with the FPA data, which only includes 753 sectors. This leaves us with 100,000 observations and 25,000 unique firms.

To prepare the data for plotting and regression analysis at the sector level, we first exclude filings that do not mention any words related to a specific topic. Alternatively, we explore a more restrictive definition of attentiveness by discarding filing-topic combinations that mention one or fewer words on a given topic. To aggregate firms into NAICS sectors, we collapse the firm-level time series by calculating simple means of observations within each firm. We then compute the sales-weighted average of firms within each sector. This aggregation is necessary, as we do not have time series data for the FPA. The final dataset comprises a panel of sector-topic observations.