Rationally Inattentive Seller: The Geographic Dispersion of U.S. Retail Chains and its Aggregate Implications

Edson Wu*
The University of Texas at Austin

This Draft: October 29, 2024

Click for the latest version

Abstract

This paper studies how the organizational structure–whether decisions are made at headquarters or regional divisions–of multi-region firms influences their expectations formation and price-setting behavior when paying attention to shocks is costly. To do so, I develop a dynamic general equilibrium model with *multi-region, rationally inattentive* firms in which firms collect information on both aggregate and region-specific shocks. When decisions are made at headquarters, firms allocate attention between overall demand and regional demand differences, ignoring the latter as geographic dispersion increases. In contrast, when decisions are decentralized, regional divisions focus solely on their own demand. I calibrate the model to U.S. Federal Reserve districts, matching the average within-firm across-regions relative price dispersion in NielsenIQ scanner data. In the calibrated model, monetary shocks have real effects that are six times larger under regional decision-making compared to headquarters, and region-specific shocks spill over to other regions when decisions are centralized. Empirically, scanner data reveals that, even after accounting for distance, product-level relative price dispersion between regions is smaller within the same retail chain than across different chains, a result qualitatively replicated by the model.

*I am deeply grateful to Hassan Afrouzi, Saroj Bhattarai, and Olivier Coibion for their invaluable guidance and support. I also thank Luis Carvalho, Henrique Castro-Pires, Arpita Chatterjee, Andres Drenik, Chad Fulton, Benjamin Johannsen, Callum Jones, Lauri Kytömaa, Oliver Pfäuti, Daniel Villar, Choongryul Yang and seminar participants at the Federal Reserve Board, and UT Austin for their insightful comments and suggestions. Email: edsonwu@utexas.edu. Researcher(s) own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1 Introduction

Consider a grocery store, part of a local chain, and its pricing decision. The Federal Reserve is set to announce a major interest rate decision today, but it's unlikely the local chain is paying much attention to it. Its focus is on understanding its own local demand and responding to what's happening locally, so broader economic changes might not play a big role in its pricing decisions.

Now think about Walmart, which is a national chain that operates across multiple states. Unlike the local chain, Walmart needs to account not only for the local demand but also for demand across all the regions it serves. With vast information coming from various locations, tracking every local detail can be overwhelming and costly. Instead of focusing on subtle differences in local conditions, the chain may be better off by simplifying its approach, paying more attention to national trends, such as the Federal Reserve's policy decisions, which affect all their regions where it operates. A simplified approach can lead to similar pricing across the chain's locations, *even* when local economic conditions vary.

In this context, three key research questions arise: (1) How do firms operating in multiple regions manage the complexity of setting prices when paying attention to both monetary and region-specific shocks is costly? (2) How does a firm's organizational structure – whether decisions are centralized at headquarters or decentralized to regional divisions – affect its optimal information acquisition and pricing strategy? (3) How do these factors affect the real effects of monetary policy and the propagation of regional shocks across the economy?

To address these questions, this paper develops a dynamic general equilibrium model with *multi-region, rationally inattentive* firms. Specifically, firms face limited capacity to process information on the numerous shocks affecting the economy and must *optimally* decide how many signals to acquire, which linear combination of shocks the signals capture, and their precision. Once these firm features are introduced, the role of the organizational structure in decision-making becomes essential, as solving for the firm's optimal decisions as a whole is not equivalent to solving each regional division's problem individually. They also endogenously generate a relationship between the number of regions in which a firm operates and its expectations about aggregate variables, which affects its within-firm price dispersion across regions, potentially resulting in *uniform pricing*. I calibrate the model using U.S. grocery scanner data. The results show that when pricing decisions are made at the regional division level, monetary policy has about six times larger effects on aggregate GDP than when decisions are centralized at the headquarters. Additionally, under headquarters level decision-making, region-specific shocks *spill over* to other regions through the firm's network of establishments, a feature absent in regional division decision-making. Empirically, using U.S. scanner prices, I find that product-level relative price dispersion between regions increases with distance, consistent with the literature on the failure of the Law of One Price. However, even after accounting for distance, product-level relative price dispersion between regions is smaller within the same retail chain than across different chains, a result my model qualitatively replicates.

The basic model in Section 2 provides closed-form characterization of multi-region firms' optimal beliefs and optimal pricing under rational inattention in a static framework. A multi-region rationally inattentive firm must acquire information about region-specific and aggregate shocks that affect its prices, subject to a information capacity constraint. As the number of regions in which a firm operates increases, the number of fundamental shocks that directly affect firms' total profits increases as well. Firms' optimal beliefs depend on whether firms' decisions are made at the headquarters level, in a centralized way, or at the regional division level, in a decentralized way. A common feature of the firms' optimal beliefs is that they are *rationally confused* about the region-specific and aggregate shocks. In my baseline framework, I assume that decisions are made at the headquarters level. Then, I explore the case where the firm solves each region's problem independently, subject to a fraction of the firms total capacity, to investigate how a firm's organizational structure affects the propagation of shocks throughout the economy.

When decisions are made at the headquarters level, multi-region rationally inattentive firms' optimal beliefs consist of one signal about their *overall demand* condition and signals about its regions' *relative demand* conditions. When the firm is sufficiently geographically dispersed, as measured by the number of regions in which it is present, it optimally chooses to ignore the signals about its regions' *relative demand* and acquire only the signal about its overall demand condition. Within my setting, this results in *uniform pricing*–that is, firms set the same price across regions where they operate, as the firm will use only the signal about its overall demand to set prices.

This centralized decision-making leads to a dampened price response to aggregate shocks compared to the full information case[1]. Firms observe signals that combine both aggregate and region-specific shocks, and when an aggregate shock occurs, they are unable to fully distinguish whether it is an aggregate or a region-specific shock, resulting in a dampened price response to aggregate shocks. Additionally, a firm's price in one region is affected not only by its own region-specific shock but also by shocks in other regions where the firm operates. This creates regional

[1]This effect on prices is present even when a firm operates in a single region and it is akin to Lucas (1972) signal-extraction problem.

spillovers of these shocks. Since firms receive noisy signals about their overall and regions' relative demand across regions, they are unable to fully identify the origin of a region-specific shock. As a result, prices in one region are affected by shocks in other regions where the firm operates.

When decisions are decentralized to regional divisions, each division operates as if it were a single-region rationally inattentive firm, with access to only a fraction of the firm's total capacity. Each division's optimal belief consists of a single signal about its *own demand* conditions. Like headquarters level decision-making, this leads to dampened price responses to aggregate shocks compared to the full information case. However, unlike in the headquarters level decision-making, a region's price is only affected by its own region's shocks. As a result, there are no regional spillovers of shocks through the firm's internal network of establishments.

To assess the macroeconomics impacts of organizational structure in multi-region, rationally inattentive firms, Section 3 extends the basic static framework into a dynamic general equilibrium model to compare the propagation of monetary and region-specific shocks when decisions are made at the headquarters versus the regional division level. The economy consists of a finite number regions, each with a representative household that has preferences over different retailers. The model features a common monetary shock, and region-specific markup shocks. Retailers are heterogeneous in the set of regions they operate in, which is given *exogenously*, disciplined using the NielsenIQ scanner data, and set prices under monopolistic competition within those regions. Importantly, they are *rationally inattentive* and take decisions at the headquarters level. Given the constraints on the amount of information (capacity) they can process about the numerous shocks affecting the economy, firms must decide how many signals to acquire, what linear combination of shocks these signals should capture, and the precision with which they observe them.

The quantitative model is calibrated using data from the United States, where each region corresponds to one of the twelve Federal Reserve districts, assuming that decision-making is made at the headquarters level. The supply side of the economy is calibrated using NielsenIQ scanner data from U.S. grocery retailers. A key contribution of this study is the calibration of firms' information capacity, using a specific moment from the data that reflects how informationally constrained firms are. When a firm has limited capacity, its within-firm price dispersion across markets remains small relative to the economy-wide price dispersion, regardless of how geographically dispersed the firm is. As the firm's capacity increases, it begins to set prices closer to what would occur under full information, leading to greater within-firm price dispersion across markets relative to the economy-wide price dispersion. Averaging this relative within-firm price dispersion across firms and over time reveals a monotonic relationship between price dispersion and firms' capacity, enabling my calibration strategy.

Once I calibrate the model, I compare the cumulative impulse response of aggregate GDP to a expansionary monetary shock under headquarters against regional division level decision-making. I find that under regional division level decision-making, monetary shocks have six times larger real effects on aggregate GDP compared to headquarters level decision-making. On the one hand, when decisions are made at the headquarters level, the firm has to divide its capacity in potentially many signals, which lowers the capacity dedicated to the overall demand signal. On the other hand, when decisions are made at the regional division level, while each regional division puts all of its capacity into the signal about its own demand, its capacity is just a share of firm's total capacity, as firm basically reallocates its total capacity across its regional divisions.

The quantitative implications of organizational structure for the propagation of region-specific shocks depend on the specific region affected. Although the aggregate effects of regional shocks are not significantly different between the two organizational structures – with the cumulative impulse response of aggregate GDP to a positive regional markup shock decline being between 0.5% and 10% lower under regional divisions compared to headquarters decision-making – the *distributional* effects are substantial. In an economy where regional divisions make decisions, the markup shock remains confined to the affected region. However, in the calibrated model, where decisions are made at the headquarters level, the firm's network of locations allows regional shocks to propagate across multiple regions. For instance, a positive markup shock in the Atlanta Fed district causes a contraction in the district's GDP, both in the calibrated economy and in the economy with regional division decision-making. However, under regional division decision-making, the shock is contained within the Atlanta Fed district and does not spill over to other regions. In contrast, in the calibrated economy, the shock propagates through the firm's network of locations, affecting all regions except the Boston and New York Fed districts. Moreover, the spillover is quantitatively significant for certain regions; for example, the GDP contraction in the Cleveland Fed district amounts to about 15% of the contraction experienced in the Atlanta Fed district, where the shock originated.

Finally, I validate the model by examining gravity-type regressions, similar to Engel and Rogers (1996), and assessing whether relative-price dispersion between counties decreases when comparing prices within the same chain. I first find that the relative-price dispersion of the same product between counties is smaller when you compare prices for the same chain, even after controlling for distance. Using NielsenIQ Scanner Data, I construct a dataset with random pairs of prices for the same good, allowing for variation in location and firm. This enables comparisons of prices both across regions and firms. For each pair, I calculate the standard deviation of their log price difference over time. Cross-sectional regressions confirm that price dispersion decreases when comparing prices within the same chain, even accounting for county distance. When I replicate this regression within my calibrated model, the results shows qualitatively that price comparisons within the same chain reduce relative-price dispersion between regions, giving support that rational inattention can be an explanation for the low within-firm price dispersion across markets that we see in the data.

Related literature. This paper is motivated by a literature that documents that retail chains charge nearly uniform pricing across markets. DellaVigna and Gentzkow (2019) documents uniform pricing in U.S. grocery goods and suggests managerial inertia[2] as a likely mechanism for this. Similarly, Daruich and Kozlowski (2023) finds evidence of uniform pricing in grocery prices in Argentina, showing that this affects the elasticity of prices to regional versus aggregate shocks. Garcia-Lembergman (2020) shows that county-level prices are affected by local demand shocks from other counties served by the same retail chains, suggesting that uniform pricing is the underlying reason. This phenomenon extends beyond prices, as Hazell, Patterson, Sarsons, and Taska (2023) studies it in the context of national wage setting, and Hyun and Kim (2019) in the context of uniform product replacement. These papers take uniformity of firms' decisions across establishments as a constraint for them and then explore the implications of the optimizing behavior under this assumption. In contrast to them, I provide a theory that can *endogenously* generate uniformity of actions across establishments, particularly focusing on the limited variation of within-firm prices across markets.

This paper also speaks to the literature on spatial spillovers of local shocks through a firm's internal network. For instance, Giroud and Mueller (2019) explore how financially constrained, multi-region firms propagate local shocks through internal resource allocation, while Gumpert, Steimer, and Antoni (2021) highlight that the managerial organization of multiestablishment firms is interdependent across establishments. My model adds to this literature by focusing on firms' optimal information acquisition about the different shocks in the economy as a story that can generate regional spillovers of local shocks.

Furthermore, my work is also related to the pricing-to-market literature that explores why the law of one price does not hold. Using the NielsenIQ scanner data, I run gravity type regressions as in Engel and Rogers (1996) and Broda and Weinstein (2008) to show that a product's relative price dispersion between counties decreases when we compare prices within the same chain, as

[2]Which includes both agency frictions and behavioral factors that prevents firms from setting optimal prices.

opposed to prices between chain, even after controlling for distance. My findings bridge the literature on why the law of one price does not hold with the fact that within-firm price dispersion across markets is low. On the theory side, I provide a novel mechanism that can explain this fact.

Building on these empirical insights, I develop a model that builds on the rational inattention in monetary economics literature. This literature proposes that costs of information processing can lead to inertia in price responses to monetary shocks observed in the data and large monetary non-neutrality (Sims, 2003, 2010; Mackowiak and Wiederholt, 2009). Matějka (2015) and Stevens (2019) show that rational inattention can endgenously generate discreteness in prices in the time series, a fact that is observed in the data. My paper adds to this literature by showing that rational inattention can also lead to price coarseness across establishments within a firm.

To explain this price coarseness, I draw on rational inattention models that incorporate multiple actions and multiple states. While Pasten and Schoenle (2016) and Yang (2022) examine multi-product rationally inattentive firms and their implications for monetary non-neutrality under exogenous signal loadings, I instead solve for a general signal structure. Relatedly, Fulton (2022) uses a two-market firm pricing problem to show how information acquisition costs can lead to uniform pricing. My approach expands on this by considering an arbitrary number of regions, illustrating that the *number of regions* in which a firm operates can contribute to uniform pricing. Additionally, I build a dynamic general equilibrium model to assess the implications of a firm's organization structure for monetary non-neutrality.

Finally, this paper builds on the literature studying dynamic rational inattention problems (DRIP) in linear-quadratic gaussian settings (Afrouzi and Yang, 2021; Miao, Wu, and Young, 2022). In particular, I use the toolkit developed in Afrouzi and Yang (2021) to solve for the optimal steady-state information structure.

2 Static Pricing for a Rationally Inattentive Multi-Region Firm

In this section, I study how the number of regions in which a firm operates affects its expectations and pricing behavior. I begin by considering a static framework where a rationally inattentive, multi-region firm determines how much information to acquire and how to set its prices. The only source of heterogeneity in this framework is the *number of regions* a firm operates. This is exogenous in my framework. The model offers closed-form solutions, providing insight into when the firm opts for a single signal about its overall demand – a *uniform pricing* policy – versus acquiring as many signals as regions where it operates – a *pricing-to-market* policy. Finally, I explore the case where the firm solves each region's problem independently to investigate how a firm's organizational structure affects the propagation of shocks throughout the economy.

2.1 Environment

Consider an economy with $n \in \mathbb{N}$ *symmetric* locations, each island indexed by $l = 1, 2, \ldots, n$. Consider a firm j that is present in k islands, $k \in \mathbb{N}, k \le n$. Let $[k]$ be the set of locations in which the firm is present. Given the symmetry of the locations, assume, without loss of generality, that $[k] = \{1, 2, \ldots, k\}$. There are $n+1$ fundamental shocks: one common monetary shock, m, and n island-specific shocks, $\{\lambda_l\}_{l\in[n]}$,such that the vector of shocks is given by $\vec{x} = (m, \lambda_1, \ldots, \lambda_n)' \in \mathcal{N}(\vec{0}, \mathbf{I}_{n+1})$. Let $\pi_j(\vec{p}_j; \vec{x})$ be the profit of the firm when it charges $\vec{p}_j = (p_{1j}, p_{2j}, \ldots, p_{kj})' \in \mathbb{R}^k$. If the firm was able to perfectly observe all shocks and in the absence of any friction, it would choose

$$\vec{p}_j^{\diamond} \equiv \max_{\vec{y}\in\mathbb{R}^k} \pi_j(\vec{y}; \vec{x})$$

where $\vec{p}_j^{\diamond}$ are firm's ideal prices. If it does not set prices equal to their ideal levels, it incurs quadratic *total* profit losses, which are given by

$$L(\vec{p}_j; m, \lambda_1, \ldots, \lambda_n) = \sum_{l\in\{1,\ldots,k\}} \frac{B}{2}(p_{lj} - p_{lj}^{\diamond})^2$$

with $\frac{B}{2}(p_{lj} - p_{lj}^{\diamond})^2$ being the region l's profit loss. B is a concavity parameter of the profit loss function[3] and $p_{lj}^{\diamond}$ is firm's j ideal price in island l, defined as the price that firm j would set in location l in the absence of any friction, and it is given by

$$p_{lj}^{\diamond} = m + \lambda_l, \quad l \in \{1, \ldots, k\} \quad \text{or in matrix form} \quad \vec{p}_j^{\diamond} = \mathbf{H}_k' \vec{x}$$

where $\mathbf{H}_k \in \mathbb{R}^{(n+1)\times k}$ is a matrix that maps the fundamental shocks $\vec{x}$ into the ideal prices and it is indexed by k to denote that it is a the matrix corresponding to a firm that is in $[k]$. While here I assume that the ideal price has this form, it arises endogenously in my dynamic framework. In the absence of any friction, firm j chooses $p_{lj} = p_{lj}^{\diamond}, \forall l \in [k]$, and the profit losses are zero. However, firms are rationally inattentive.

[3] In principle, this may depend on l and j. However, as a result of the assumption of symmetric islands, $B_{lj} = B$, $\forall l \in [k], \forall j$. In the quantitative model, this will be a function of l and the number of regions a firm is present.

2.2 Economics of Attention Allocation

Assume that firm j is subject to a finite attention capacity $\kappa \in \mathbb{R}_+$ and must optimally allocate this capacity into different signals – noisy linear combinations of the shocks. Given the chosen signals, the firm chooses a price function that maps these signals into actions.

Optimal Information Allocation When Decisions Are at the Headquarters Level. When decisions are made at the headquarters level, the headquarters optimally chooses signals and the price function considering the firm's *total* profit loss. The attention allocation problem can be written as

$$\max_{S_{j0}\subset\mathbb{S}} \mathbb{E}\Big[\max_{\vec{p}_j:S_j^0\to\mathbb{R}^k} \mathbb{E}\big[-\sum_{l\in\{1,\dots,k\}} \frac{B}{2}(p_{lj}-p_{lj}^{\diamond})^2|S_j^0\big]\Big] \tag{1}$$

$$\text{s.t. } \mathscr{I}(S_j^0;\vec{x}|S_j^{-1}) \leq \kappa \tag{2}$$

$$S_j^0 = S_{j0}\cup S_j^{-1} \tag{3}$$

$$S_j^{-1} \text{ given} \tag{4}$$

where constraint (2) is the capacity constraint and $\mathscr{I}(S_j^0;\vec{x}|S_j^{-1})$ is the conditional Shannon's mutual information function and measures the amount of information that firm's signal history S_j^0 contains about $\vec{x}$, conditional on S_j^{-1}. The more informative is S_j^0 about $\vec{x}$ conditional on S_j^{-1}, the larger is this measure. Equation (3) captures the evolution of firm's information set. In particular, it states that the firm does not forget information. Equation (4) is the initial prior uncertainty, exogenously given. Firms have a given capacity κ and allocate it across different signals.

Equation (1) can be rewritten as a maximization problem in which the firm chooses an optimal posterior variance-covariance matrix of the shocks (Kőszegi and Matějka, 2020; Afrouzi and Yang, 2021; Fulton, 2022) and given by

$$\max_{\Sigma_0} -\frac{1}{2} tr(\Sigma_0\tilde{\tilde{\Omega}}) \tag{5}$$

$$\text{s.t. } \frac{1}{2}\ln\left(\frac{|\Sigma_{-1}|}{|\Sigma_0|}\right) \leq \kappa \tag{6}$$

$$\Sigma_{-1} - \Sigma_0 \succeq 0 \tag{7}$$

$$0 \prec \Sigma_{-1} \preceq \infty \tag{8}$$

where I omit the index j. $tr(.)$ is the trace operator, $|.|$ is the determinant operator, $\succeq$ denotes positive semidefiniteness, Σ_0 is the posterior variance-covariance matrix of the fundamental shocks given the history of signals S^0. That is, $\Sigma_0 \equiv \mathbb{E}\left[(\vec{x}-\mathbb{E}[\vec{x}|S^0])(\vec{x}-\mathbb{E}[\vec{x}|S^0])'|S^0\right]$.

$\Sigma_{-1} \equiv \mathbb{E}\left[(\vec{x} - \mathbb{E}[\vec{x}|S^{-1}])(\vec{x} - \mathbb{E}[\vec{x}|S^{-1}])'|S^{-1}\right]$ is the prior variance-covariance matrix of the fundamental shocks given S^{-1}. $\tilde{\tilde{\Omega}} \equiv \mathbf{H}\text{diag}(B)\mathbf{H}'$ is a benefit matrix that governs how important are the many fundamental shocks for the firm's profit loss, where $\text{diag}(B)$ is a matrix whose diagonal is given by B, and off-diagonal elements are zero. Constraint (6) is the capacity constraint. Finally, constraint (7) is the set of no-forgetting constraints. It states that the firm cannot choose to be more uncertain along any dimension than its prior uncertainty.

Assumption 1. $\Sigma_{-1} = \mathbf{I}$

Assumption 1 states that the firm has the same prior uncertainty regarding any shocks. In Proposition A.3, I allow prior uncertainty about the aggregate shock and regional shocks to be different.

Proposition 1. (Optimal Signal Structure under Headquarters' Pricing) Under Assumption 1, the solution to the problem in Equation (5)-Equation (8) is such that there exists a threshold $\underline{k}(\kappa) = e^{2\kappa} - 1$ such that if $k \geq \underline{k}$, the firm acquires *one* signal[4], while if $k < \underline{k}(\kappa)$, the firm acquires k signals. When the firm acquires *one* signal, it acquires a signal about its *overall demand*:

$$s_1^{\text{HQ}} = \frac{1}{\sqrt{1+\frac{1}{k}}}\Big(\underbrace{\frac{1}{k}\sum_{l=1}^{k} p_l^{\diamond}}_{\substack{\text{average}\\ \text{ideal price}}}\Big) + \nu_1^{\text{HQ}}, \quad \nu_1^{\text{HQ}} \sim \mathcal{N}\Big(0, \frac{1}{e^{2\kappa}-1}\Big)$$

When the firm acquires k signals, it acquires a signal about its *overall demand*:

$$s_1^{\text{HQ}} = \frac{1}{\sqrt{1+\frac{1}{k}}}\Big(\underbrace{\frac{1}{k}\sum_{l=1}^{k} p_l^{\diamond}}_{\substack{\text{average}\\ \text{ideal price}}}\Big) + \nu_1^{\text{HQ}}, \quad \nu_1^{\text{HQ}} \sim \mathcal{N}\left(0, \frac{1}{(k+1)^{1-\frac{1}{k}}(e^{2\kappa})^{\frac{1}{k}} - 1}\right)$$

and $k-1$ signals about regional *relative demands*:

$$s_i^{\text{HQ}} = \frac{1}{\sqrt{1+\frac{1}{i-1}}}(p_{i-1}^{\diamond} - \frac{1}{i-1}\sum_{m=2}^{i-1} p_{m-1}^{\diamond} - \frac{1}{i-1}p_k^{\diamond}) + \nu_i^{\text{HQ}}, \quad \nu_i^{\text{HQ}} \sim \mathcal{N}\left(0, \frac{1}{(k+1)^{-\frac{1}{k}}(e^{2\kappa})^{\frac{1}{k}} - 1}\right) \quad i \in \{2, \ldots, k\}$$

Proposition 1 shows the main result on how the number of regions a firm operates in affects its information acquisition. Given $\kappa > 0$, if the number of region in which a firm operates is sufficiently high, $k \geq \underline{k}(\kappa)$, the firm decides to acquire only information about its overall demand, ig-

[4] I solve for the optimal posterior, then I choose signals that generates this posterior. In principle, there is a set of signals that generates the same posterior. I choose signals such that rational inattention errors are independent, following Afrouzi and Yang (2021).

noring signals that are informative about regional relative demands. Therefore, the firm uses only this single signal to price all regions where it operates. Given the assumption on how ideal prices depend on the shocks, this leads to *uniform pricing*. If the firm does not operate in many regions, $k < \underline{k}(\kappa)$, then the firm acquires as many signals as number of regions in which it operates, leading to *pricing-to-market*. Proposition A.1 provides closed-form solutions for optimal prices under the optimal information structure in Proposition 1. With that, we can show how this mechanism affects the *within-firm, between-region* relative price dispersion and the responsiveness of the firm's prices to shocks.

Proposition 2. (Within-Firm Between-Region Relative Price Dispersion) Given a firm in k regions, the relative price dispersion between region's one and ℓ is given by

$$var(p_{1j}^{\text{HQ}} - p_{\ell j}^{\text{HQ}}) = \begin{cases} 0 & , k \geq e^{2\kappa} - 1 \\ 2(1 - ((k+1)e^{-2\kappa})^{1/k}) & , k < e^{2\kappa} - 1 \end{cases}$$

where p_{lj}^{HQ} is the optimal price of firm j in location l. Furthermore, the within-firm between-region relative price dispersion is lower than when the firm sets prices under full information. That is,

$$var(p_{1j}^{\text{HQ}} - p_{\ell j}^{\text{HQ}}) \leq var(p_{1j}^{\diamond} - p_{\ell j}^{\diamond})$$

Finally, when $k < e^{2\kappa} - 1$

$$\frac{\partial var(p_{1j}^{\text{HQ}} - p_{\ell j}^{\text{HQ}})}{\partial k} < 0, \quad \forall k \geq 2$$

That is, the relative price dispersion is decreasing in k.

Proposition 2 shows that when multi-region firms are rationally inattentive and solve its problem in a centralized way, its between-region relative price dispersion is lower than its full-information counterpart. Moreover, it decreases with the number of regions in which the firm operates. When the number of region in which the firm operates is sufficiently high, the firm acquires a single signal and the between-region relative price dispersion becomes zero.

Proposition 3. (Firm's Price Response to Monetary Shocks under Headquarters' Pricing) Firm j's price response to an expansionary monetary shock is given by:

$$\frac{\partial p_{lj}^{\mathrm{HQ}}}{\partial m} = \begin{cases} (1 - e^{-2\kappa}) & , k > \underline{k}(\kappa) \\ 1 - \frac{1}{(k+1)^{1-\frac{1}{k}} (e^{2\kappa})^{\frac{1}{k}}} & , k \leq \underline{k}(\kappa) \end{cases}$$

Furthermore, there's $\underline{k}^* < \underline{k}(\kappa)$ such that

$$\frac{\partial\left(\partial p_{lj}^{\mathrm{HQ}} / \partial m\right)}{\partial k} = \begin{cases} \geq 0 & , k \geq \underline{k}^* \\ < 0 & , k < \underline{k}^* \end{cases}$$

That is, a firm's regional price response to the common shock is *non-monotonic* in k. This reflects two channels: 1) Increasing k adds more signals to which firms want to pay attention to and firms reallocate capacity from all other signals to this new signal; 2) Increasing k increases the relative importance of the signal about the overall demand conditions.

Proposition 3 directly speaks to how monetary shocks have real effects in the economy. The more responsive prices are to monetary shocks, the lower is monetary non-neutrality. As a result, the extent to which increasing k increases or decreases monetary non-neutrality depends on whether k is below or above $\underline{k}^*$[5].

Proposition 4. (Firm's Price Response to Island-Specific Markup Shocks under Headquarters' Pricing) Firm j's price response in island $l \in [k]$ to a positive markup shock in island ℓ is given by:

1. When $l = \ell$,

$$\frac{\partial p_{lj}^{\mathrm{HQ}}}{\partial \lambda_\ell} = \begin{cases} (1 - e^{-2\kappa})\frac{1}{k} & , k \geq e^{2\kappa} - 1 \\ 1 - \frac{k}{(k+1)^{1-\frac{1}{k}} (e^{2\kappa})^{\frac{1}{k}}} & , k < e^{2\kappa} - 1 \end{cases}$$

   In this case, the price response in the island where the shock occurs is *dampened* compared to the full information case.

[5]The non monotonicity in Proposition 3 is also present in Pasten and Schoenle (2016) in a setting where signal loadings are *exogenous* and are signals about the fundamental shocks plus noise. In their setting, a firm in k regions receives $k+1$ signals: one signal about the aggregate shock, and k signals about regional shocks. Each signal is of the form of true value plus noise.

2. When $l \neq \ell$,

$$\frac{\partial p_{lj}^{\text{HQ}}}{\partial \lambda_\ell} = \begin{cases} (1-e^{-2\kappa})\frac{1}{k} & ,k \geq e^{2\kappa}-1 \\ \frac{1}{(k+1)^{1-\frac{1}{k}}(e^{2\kappa})^{\frac{1}{k}}} & ,k < e^{2\kappa}-1 \end{cases}$$

In this case, markup shocks in island ℓ *spill over* to other islands served by the firm. Under full information, prices in islands not directly affected by the markup shock *do not* respond.

Proposition 4 shows how rational inattention can generate a new channel for propagating regional shocks across space through multi-region firms' information acquisition and their network of locations. In the extreme case, it can *endogenously* lead to uniform pricing at the firm level, even when they can choose different prices for each location. Even when firms do not adopt uniform pricing, the mechanism still generates regional spillovers of local shocks.

Now, I move to the case in which decisions are made at the regional division level to explore how a firm's organization structure affect information acquisition and pricing. The following results will show that when information acquisition is costly, it is important to consider how decisions are made within a firm.

Optimal Information Allocation When Decisions Are at the Regional Division Level. In the previous analysis, the firm was assumed to solve the information acquisition and pricing problems centrally at the headquarters level. In this subsection, I shift the focus to a decentralized approach, where firm j solves the problems for each region individually and independently. Each regional division l operates under a capacity constraint that represents a fraction of the firm's total capacity and seeks to minimize its *own* region's profit loss. Thus, the attention allocation problem for regional division l of firm j can be formulated as follows

$$\max_{S_{lj0}\subset\mathbb{S}} \mathbb{E}\big[\max_{p_{lj}:S_{lj}^0\to\mathbb{R}} \mathbb{E}\big[-\frac{B}{2}(p_{lj}-p_{lj}^{\diamond})^2|S_{lj}^0\big]\big] \tag{9}$$

$$\text{s.t. } \mathscr{I}(S_{lj}^0;\vec{x}|S_{lj}^{-1}) \leq \kappa_l \tag{10}$$

$$S_{lj}^0 = S_{lj0} \cup S_{lj}^{-1} \tag{11}$$

$$S_{lj}^{-1} \text{ given} \tag{12}$$

where signals are now indexed by l, reflecting the fact that each region chooses a signal structure. Furthermore, each region chooses only its island prices. Therefore, the pricing function is a map from signals to $\mathbb{R}$. We can write the linear-quadratic gaussian problem associated with it.

$$\max_{\Sigma_{l,0}} -\frac{1}{2} tr(\Sigma_{l,0}\tilde{\tilde{\Omega}}_l) \tag{13}$$

$$\text{s.t. } \frac{1}{2}\ln\left(\frac{|\Sigma_{l,-1}|}{|\Sigma_{l,0}|}\right) \le \kappa_l \tag{14}$$

$$\Sigma_{l,-1} - \Sigma_{l,0} \succeq 0 \tag{15}$$

$$0 \prec \Sigma_{l,-1} \preceq \infty \tag{16}$$

where I omit the index j. All objects are now indexed by l, as each l solves a different problem. Importantly, $\tilde{\tilde{\Omega}}_l \equiv (\mathbf{e}_1 + \mathbf{e}_{l+1})B(\mathbf{e}_1 + \mathbf{e}_{l+1})' \in \mathbb{R}^{(n+1)\times(n+1)}$ depends on l.

Assumption 2. $\Sigma_{l,-1} = \mathbf{I}$

Assumption 2 states that the regional division l has the same prior uncertainty regarding any shocks.

Assumption 3. *Assume that for a firm that is in k islands, the firm divides its capacity across its regions such that* $\sum_{l\in[k]} \kappa_l = \kappa$, $\kappa_l \ge 0$.

Proposition 5. (Optimal Signal Structure under Regional Division Pricing) Under Assumption 2 and Assumption 3, the solution to the problem in Equation (13)-Equation (16) is such the firm acquires *one* signal about its *own demand*:

$$s_l^{\text{RD}} = \frac{1}{\sqrt{2}} \underbrace{p_l^{\diamond}}_{\substack{l\text{'s own} \\ \text{ideal price}}} + v_l^{\text{RD}}, \quad v_l^{\text{RD}} \sim \mathcal{N}\left(0, \frac{1}{e^{2\kappa_l} - 1}\right) \tag{17}$$

There is a clear distinguishing feature of signal s_l^{RD} in Proposition 5 and the signal about overall demand s_1^{HQ} in Proposition 1. While s_1^{HQ} depends on *all* location-specific shocks, s_l^{RD} depends only on the location's l shock. As a result, when firms set prices at the regional division level, there is no regional spillovers of island-specific shocks. When firms set prices at the headquarters level, island-specific shocks spill over across regions served by the firm.

Proposition 6. (Firm's Price Response to Monetary Shocks under Regional Division Pricing) Firm j's price response in region l to an expansionary monetary shock is given by:

$$\frac{\partial p_{lj}^{\text{RD}}}{\partial m} = (1 - e^{-2\kappa_l})$$

If we further assume that $\kappa_l = \frac{\kappa}{k}$, that is, κ is divided equally across the regions where the firm in k

islands operates

$$\frac{\partial\left(\partial p_{lj}^{\mathrm{RD}}/\partial m\right)}{\partial k} < 0$$

When $\kappa_l = \frac{\kappa}{k}$, we can go further and show that when the firm solves each region's problem in a decentralized way, its price response to monetary shocks in *every* region where it operates is dampened compared to the pricing in a centralized way. With that we can establish Corollary 1.

Corollary 1. Consider the environment in Section 2.1 and problems in Equation (5)-Equation (8) and Equation (13)-Equation (16). Consider a firm in k islands. Assume that $\kappa_l = \kappa/k$. Then, as established in Propositions 3 and 6, we have that a firm's price in any given island $l \in [k]$ responds more to monetary shocks under headquarters pricing as it does under regional division pricing.

Corollary 1 shows a key result to understand how monetary shocks affects the real economy under headquarters and regional division pricing. A firm's price respond more to monetary shocks under headquarter pricing than under regional division pricing. As a result, monetary shocks will have lower real effects on the economy – that is, monetary non-neutrality decreases. Under headquarters decision-making, the firm allocates its capacity across multiple signals, placing greater weight on the signal about overall demand. In contrast, under regional division decision-making, the firm concentrates its capacity on a single signal, but this capacity represents only a fraction of what is available when decisions are made at the headquarters.

Proposition 7. (Firm's Price Response to Island-Specific Shocks under Regional Division Pricing) Firm j's price response in island $l \in [k]$ to a positive markup shock in island $\ell \in [k]$ is given by

$$\frac{\partial p_l^{\mathrm{RD}}}{\partial \lambda_\ell} = \begin{cases} (1 - e^{-2\kappa_l}) & , \ell = l \\ 0 & , \ell \neq l \end{cases}$$

and markup shocks in island ℓ *do not* spill over to other islands served by the firm.

When decision-making occurs at the regional division level, the price in a given region $l \in [k]$ does not respond to region-specific markup shocks in other regions where the firm operates. This is because regional divisions focus solely on their own region's profit losses and therefore acquire signals about their own region's demand. In contrast, under headquarters-level decision-making, a markup shock in one region prompts the firm to adjust its prices across other regions, as the firm relies on noisy signals that prevent it from fully identifying the source of the shock. With that, we can establish Corollary 2.

Corollary 2. Consider the environment in Section 2.1 and problems in Equation (5)-Equation (8) and Equation (13)-Equation (16). Then, as established in Propositions 4 and 7, when firms set prices at the headquarters level, island-specific markup shocks in regions where the firm operates *spill over* to firm's prices in other regions, while when firms set prices at the regional division level, there are no spillovers.

Although it may be challenging to directly observe how firms make decisions, Corollary 2 offers a testable implication for determining whether firms are making decisions at the regional or headquarters level when information acquisition is costly. If decisions are made at the regional division level, we should see no spillovers of local shocks across the firm's network of locations. Conversely, if decisions are centralized at the headquarters, we should[6].

Extensions. While I derive Proposition 1 under the assumption of a fixed capacity κ that does not vary with k, Proposition A.2 shows similar result under conditions in which κ increases with k. In Proposition A.3, I consider the case in which the firm has a different prior regarding the aggregate shock and the regional shocks. In this case, the threshold at which firms start acquiring only one signal depends also on the relative prior variances regarding aggregate and regional shocks.

Discussion on how multi-region firms set prices. In standard multi-region models with monopolistic competitive firms, 'firms' and 'establishments' are often used interchangeably. This is because the optimal pricing decision for a firm with multiple establishments is equivalent to solving the pricing problem for each establishment individually. However, in my framework, this equivalence doesn't hold. It is important to distinguish whether pricing decisions are made at the firm level (in my model, the headquarters) or at the establishment level (in my model, the regional division).

In the dynamic general equilibrium model, I assume that information acquisition and price setting decisions are made at the headquarters level. Studies that seek to understand multi-store retailers' pricing decisions suggest this is the case for some of them. For instance, Levy, Bergen, Dutta, and Venable (1997), in a study seeking to measure menu costs for supermarket chains, mentions that *"In the supermarket chains we study, prices are generally set at corporate headquarters in a weekly meeting where the manager in charge of setting prices looks at a variety of information (...)"*. Adams and Williams (2019), studying pricing of home-improvement retailers, mentions that

[6] Garcia-Lembergman (2020) shows evidence that can be consistent with this. In contrast to Garcia-Lembergman (2020), the spillover is present even when firms are not literally setting uniform prices.

"Product category managers working at corporate headquarters–not local managers–make pricing and assortment decisions".

3 Dynamic General Equilibrium Model

Building on the insights from the static framework in Section 2, this section develops a dynamic general equilibrium model to quantify the effects of rationally inattentive, multi-regional firms on the propagation of both monetary and region-specific shocks.

3.1 Environment

Time is discrete, $t = 0, 1, \ldots$. The economy is composed by a discrete number of regions $n \in \mathbb{N}$, $l \in \{1, 2, \ldots, n\}$ integrated in a monetary and fiscal union. Each region is characterized by a representative household, with population $\zeta_l \in (0,1), \forall l \in [n], \sum_l \zeta_l = 1$. There is a measure one of monopolistically competitive retailers in the economy, and $2^n - 1$ types of retailers. Each type of retailer is present in a subset of the regions, where $2^n - 1$ is the total number of possible combinations of regions in which a firm can be present[7]. Let $\mathscr{J}_l$ be the set of firms present in l. Assume there is a measure $\varphi_h \in [0,1)$ of each type of firm, where $\sum_h \varphi_h = 1, h = 1, 2, \ldots, 2^n - 1$, and there exists $\iota(h) : \{h \in \mathbb{N} | 1 \leq h \leq 2^n - 1\} \to \{1,0\}^n \backslash (0, \ldots, 0)$. $\iota(h)$ is a function that maps indices into a firm type. The retailers use labor to produce and supply the good to the household of the region where they are present. Households save using a riskless bond. There is a single competitive labor market in this economy.

[7]A firm type can be summarized by a vector $\{0,1\}^n$. There are $2^n - 1$ possible combinations of types, where -1 excludes $\{0, \ldots, 0\} \in \mathbb{R}^n$. Within a region, the measure of firms present in it is given by 2^{n-1}. To see why, let $n = 4$, and take for instance a firm in region 1: $\{1, a, b, c\}$, where $a, b, c \in \{0,1\}$. There are 2^{4-1} possible combinations left for a, b, c.

3.1.1 Household

The representative household in region $l \in [n]$ demands a composite good and supply labor in a competitive market. The household solves

$$\max_{\{C_{lt}, L_{lt}\}_{t\geq 0}} \mathbb{E}_0^f \Big[\sum_{t=0}^{\infty} \beta^t \big(\log(C_{lt}) - L_{lt} \big) \Big] \tag{18}$$

$$\text{s.t.} \quad \int_0^1 P_{ljt} C_{ljt} dj + B_{lt} \leq W_t L_{lt} + (1 + i_{t-1}) B_{lt-1} + \text{Profits}_{lt} - T_{lt} \tag{19}$$

$$C_{lt} = \Big(\int_0^1 \theta_{lj}^{\frac{\Lambda_{lt}-1}{\Lambda_{lt}}} C_{ljt}^{\frac{1}{\Lambda_{lt}}} dj \Big)^{\Lambda_{lt}} \tag{20}$$

where $\mathbb{E}_t^f[.]$ is the full information rational expectations operator at t. C_{lt} is the aggregate consumption of household l at t, L_{lt} is her labor supply, Profits_{lt} are the profits rebated to her, and T_{lt} lump-sum transfers used to eliminate steady-state inefficiencies of monopolistic competition. The household chooses how much to spend in each retailer and how much to save in bonds subject to its total resources, given by its labor income, savings from previous period, and profits and lump-sum transfers. Note that the consumption aggregator is a CES aggregator, with θ_{lj} being the taste shifter of household l for retailer j, and it is assumed that this taste shifter is positive if j is present in l and zero otherwise. Furthermore, Λ_{lt} is the mark-up implied by the elasticity of substitution across varieties, which is allowed to vary over time and is specific to region l.

3.1.2 Retailers

Let $\mathscr{L}_j$ be the set of locations $l \in [n]$ in which retailer $j \in [0,1]$ is present. Assume that each retailer produces its output using a linear technology in labor, given by $Y_{ljt} = L_{ljt}$, where L_{ljt} is the amount of labor used by j in l to produce Y_{ljt}. Furthermore, given its demand in each location in which it is present, $l \in \mathscr{L}_j$, it chooses the price in each location in order to maximize its expected discounted sum of profits, where the per-period profit when it sets price $(P_{ljt})_{l\in\mathscr{L}_j}$ is given by

$$\Pi_{jt} = \sum_{l\in\mathscr{L}_j} \big(P_{ljt} Y_{ljt}^s - W_t L_{ljt} \big), \quad \text{s.t. } Y_{ljt}^s = \zeta_l C_{ljt}$$

recalling that C_{ljt} is the demand per capita of region l for retailer j. Therefore, to get the region l's total demand for retailer j, we need to multiply by its population.

Full information pricing. When the monopolistic competitive firm j is not subject to frictions of any kind, maximizing the joint profit in all locations is the same as maximizing the profit in each location individually. That is, for each $l \in [\mathscr{L}_j]$, the retailer j sets its price equal to its ideal price, $P^{\diamond}_{ljt}$:

$$P^{\diamond}_{ljt} \equiv \text{argmax}_{P_{ljt}} \big(P_{ljt} Y^s_{ljt} - W_t L_{ljt}\big), \quad \text{s.t. } Y^s_{ljt} = \zeta_l C_{ljt} \tag{21}$$

where the ideal price $P^{\diamond}_{ljt}$ is defined by Equation (21) as the price that firm j would set in region l and time t in the absence of any frictions.

Rational inattention. Monopolistically competitive firms are rationally inattentive. That is, they do not perfectly observe the fundamental shocks, and must optimally choose prices conditional on signals that are endogenously chosen by them, subject to a capacity constraint. That is, given a capacity κ_j, the amount of information the firm can acquire is given by

$$\mathbb{I}(S^t_j; \vec{x}^t | S^{t-1}_j) \leq \kappa_j$$

where $\mathbb{I}(S^t_j; \vec{x}^t | S^{t-1}_j)$ is the conditional Shannon mutual information between the history of signals S^t_j and the history of the fundamental shocks.

After firms make their information choices, shocks and signals are drawn, and each firm observes the realization of its signals. Then, firms choose their prices conditional on their information sets. Finally, demand is realized and firms produce to meet their demand.

More specifically, given a capacity κ_j, a firm chooses a set of signals to observe, $S_{jt} \subset \mathbb{S}^t$, and a pricing function that maps its information set to their optimal actions, $\vec{P}_{jt} : S^t_j \to \mathbb{R}^{|\mathscr{L}_j|}$, where $S^t_j = \{S_{j\tau}\}^t_{\tau=-1}$ is the firm's information set at time t and $|\mathscr{L}_j|$ is the number of regions in which firm j is present. The rationally inattentive retailer solves:

$$\max_{\{S_{j,t} \subset \mathbb{S}^t, \{P_{ljt}(S^t_j)\}_{l \in \mathscr{L}_j}\}_{t \geq 0}} \quad \mathbb{E}\Big[\sum_{t=0}^{\infty} \underbrace{\beta^t W_t^{-1}}_{\substack{\text{discount}\\ \text{factor}}} \times \Big\{ \sum_{l \in \mathscr{L}_j} \Big(\underbrace{(1-\tau_l) P_{ljt} Y^s_{ljt}}_{\substack{\text{revenue}\\ \text{in } l}} - \underbrace{W_t L_{ljt}}_{\substack{\text{production}\\ \text{cost in } l}} \Big) \Big\} \Big] \tag{22}$$

$$\text{s.t.} \quad Y^s_{ljt} = \zeta_l C_{ljt}, l \in \mathscr{L}_j \qquad \text{(demand)} \tag{23}$$

$$\mathbb{I}(S^t_j; \vec{x}^t | S^{t-1}_j) \leq \kappa_j \qquad \text{(info. processing constraint)} \tag{24}$$

$$S^t_j = S^{t-1}_j \cup S_{jt}, \quad S^{-1}_j \text{ given} \qquad \text{(evolution of information set)} \tag{25}$$

where τ_l is a constant tax to firms in location l that eliminates steady-state inefficiencies coming from monopolistic competition. From now on, I assume that $\kappa_j = \kappa, \forall j \in [0,1]$. That is, all firms have the same capacity. S^{-1}_j is an initial signal. Finally, S^t_j satisfies the no-forgetting condition, which states that firms do not forget information over time. This will put an upper bound on the amount of uncertainty the firm can choose[8].

Nominal Aggregate GDP and Aggregate Prices. Since there is no investment, the aggregate real GDP in each region l is equal to consumption $Y_{lt} = C_{lt}$. Let the aggregate real GDP C_t and the aggregate prices P_t be defined, respectively, as

$$C_t \equiv \prod_{l\in[n]} C_{lt}^{\zeta_l} \quad \text{and} \quad P_t \equiv \prod_{l\in[n]} P_{lt}^{\zeta_l}$$

Therefore, the nominal aggregate GDP M_t is given by

$$M_t \equiv P_t C_t = \prod_{l\in[n]} (P_{lt}C_{lt})^{\zeta_l}$$

3.1.3 Monetary and fiscal policy

I assume there is a single monetary authority that controls the path of *nominal aggregate GDP* $\{M_t\}_{t\geq 0}$. I also assume a single fiscal authority that levies taxes or subsidizes monopolistic competitive firms' sales in each location $l \in [n]$ at a constant rate τ_l, lump-sum transferred back to households. The government budget constraint is given by

$$\sum_{l\in[n]} \int_{j\in\mathcal{J}_l} \tau_{lt} P_{ljt} Y_{ljt}^{s} dj = \sum_{l\in[n]} T_{lt}$$

The monetary authority models the nominal GDP as an random walk process:

$$\log(M_t) = \log(M_{t-1}) + \sigma_u u_t, \quad u_t \sim N(0,1) \tag{26}$$

3.1.4 Fundamental shocks

The economy is subject to $n+1$ fundamental shocks: the nominal GDP shock from Equation (26), and n regional markup shocks $\left(\log(\Lambda_{lt})\right)_{l\in[n]}$. Besides Equation (26), we have

$$\log(\Lambda_{lt}) = \log(\Lambda_l) + \sigma_{\varepsilon_l}\varepsilon_{lt},\ \varepsilon_{lt} \sim N(0,1), \quad l \in [n] \tag{27}$$

Note that while I assume that the nominal GDP process follows a random walk, the processes for the regional markup shocks follow white noise processes.

[8]See Afrouzi and Yang (2021) for details.

Equilibrium Definition. An equilibrium consists of allocations for households and firms, monetary and fiscal policies, and prices such that: (1) given prices and policies, the allocations are optimal for households and firms, and (2) markets clear. A detailed definition is in Appendix C.

3.2 Solution Method

I consider a log-linearization of this economy around an efficient steady state, with a second-order approximation of the monopolistic competitive firm's profit loss function around this efficient steady state[9]. The derivations are in Appendix F. Going forward, small letters denote log deviations of the corresponding variables from their steady-state values[10].

Monopolistic competitive retailer's marginal cost. A firm's marginal cost in location $l \in \mathscr{L}_j$ is

$$mc_{lt} = w_t$$

Ideal prices. A firm's ideal price in location $l \in \mathscr{L}_j$ is

$$p^{\diamond}_{ljt} = \lambda_{lt} + mc_t \tag{28}$$

Even though I index $p^{\diamond}_{ljt}$ by j, note that this does not depend on j. This means that I abstract from chain-specific preference shocks or chain-specific technology shocks. Note that Equation (28) is the same as in Section 2.1.

Wages. Using the the first order conditions of the household in each location $l \in [n]$ and the monetary rule, the nominal wage in the economy is

$$w_t = m_t$$

Regional aggregate price and regional real GDP. The aggregate price in region $l \in [n]$ is

$$p_{lt} = \int_0^1 \theta_{lj} p_{ljt} dj, \quad l \in [n]$$

The regional GDP, y_{lt}, is equal to aggregate consumption, as there is no investment. Under fully elastic labor supply, this is given by the difference between the wage in location l and aggregate

[9] For a discussion on the use of second-order approximations of the objective function, see Afrouzi and Yang (2021).

[10] i.e., $x_t \equiv \log X_t - \log X$

prices in region l, p_{lt}:

$$c_{lt} = w_t - p_{lt}$$

Note that under the assumptions I made regarding household's preferences, single labor market, regional *nominal* GDP is the same for *all* regions. However, regional *real* GDP might still be different, as each region has a different set of retailers, with distinct incentives to acquire information.

Aggregate price, aggregate real GDP, and aggregate nominal GDP. The aggregate price, the aggregate real GDP, and the aggregate nominal GDP in the economy are given, respectively, by

$$p_t = \sum_{l\in[n]} \zeta_l p_{lt}, \quad c_t = \sum_{l\in[n]} \zeta_l c_{lt} \quad \text{and} \quad m_t \equiv p_t + c_t$$

Fundamental shocks. Let the vector of fundamental shocks be denoted by $\vec{x}_t \equiv (m_t, \lambda_{1t}, \ldots, \lambda_{nt})' \in \mathbb{R}^{n+1}$. Then, its state-space representation is given by

$$\vec{x}_t = \mathbf{A}\vec{x}_{t-1} + \mathbf{Q}\vec{u}_t, \quad \vec{u}_t \perp \vec{x}_{t-1}, \quad \vec{u}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

where $\vec{u}_t = (u_t, \varepsilon_{1t}, \ldots, \varepsilon_{nt})' \in \mathbb{R}^{n+1}$.

3.2.1 An approximate problem

I use a second-order approximation to the firm's problem to solve its information acquisition problem. Before turning to the rational inattention problem, it is useful to state the optimal pricing *given* a signal structure.

Imperfect information. Given a history of signals S_j^t, a firm j's optimal price in location $l \in \mathcal{L}_j$ is

$$p^*_{ljt} = \mathbb{E}[p^\diamond_{ljt} | S_j^t]$$

That is, the optimal price under imperfect information is the expected ideal price given the information set S_j^t.

Rational inattention. After performing a second-order approximation of the firm's profit function, its approximate profit maximization problem becomes

$$\max_{\{S_{jt},(p_{ljt}(S_j^t))_{l\in\mathscr{L}_j}\}_{t\geq 0}} \mathbb{E}[\sum_{t=0}^{\infty}\beta^t\{-\sum_{l\in\mathscr{L}_j}\underbrace{\tfrac{1}{2}B_{lj}(p_{ljt}(S_j^t)-p_{ljt}^{\diamond})^2}_{\text{loss from mispricing in } l}\}|S_j^{-1}] \tag{29}$$

$$\text{s.t. } \mathscr{I}(S_j^t;\{\vec{x}_\tau\}_{\tau\leq t}|S_j^{t-1})\leq\kappa_j \qquad \text{(info. processing constraint)} \tag{30}$$

$$S_j^t = S_{jt}\cup S_j^{t-1},\quad S_j^{-1}\text{ given} \qquad \text{(evolution of information set)} \tag{31}$$

where $p_{ljt}^{\diamond}=\lambda_{lt}+m_t, l\in\mathscr{L}_j$ is the price that firm j would set in location l at time t in the absence of any friction. B_{lj} is the curvature parameter of the profit function around the optimal price in a given location l for firm j. Considering the assumptions regarding the demand function,

$$B_{lj} = \underbrace{\zeta_l}_{\substack{\text{Population size}\\\text{in region } l}} \times \underbrace{\theta_{lj}}_{\substack{\text{Taste shifter in}\\\text{region } l \text{ for firm } j}} \times \frac{\Lambda_l}{\Lambda_l-1}$$

where Λ_l is the steady state markup in region l. Firm j's sales in region $l\in\mathscr{L}_j$ in the efficient steady state is

$$\mathsf{sales}_{lj} = \Lambda_l\zeta_l\theta_{lj}$$

where I consider a steady state with $M=1$, and I normalize $\int_0^1\theta_{lj}dj=1,\forall l\in[n]$. As a result, we can rewrite B_{lj} as

$$B_{lj} = \sum_{l\in\mathscr{L}_j}\Lambda_l\zeta_l\theta_{lj}\times\frac{\Lambda_l\zeta_l\theta_{lj}}{\sum_{l\in\mathscr{L}_j}\Lambda_l\zeta_l\theta_{lj}}\times\frac{1}{\Lambda_l-1} = \mathsf{sales}_j\times\mathsf{sales\ share}_{lj}\times\frac{1}{\Lambda_l-1}$$

where $\mathsf{sales\ share}_{lj}$ is region's l share of firm's j total sales. Therefore, the profit losses arising from a given mistake in pricing in region l is (1) increasing in firm's size, measured by its total sales; (2) increasing in region's l share of firm j total sales; (3) increasing in region's l elasticity of demand (which implies in lower Λ_l).

In what follows, I assume that the firm is constrained to choose gaussian signals so that we can write the problem in Equation (29) to Equation (31) as a linear-quadratic gaussian problem where we choose the optimal posterior. Furthermore, I solve for the steady-state information structure following Afrouzi and Yang (2021), where posterior, prior, benefit matrix, and shadow value of capacity constraint are time-invariant $(\bar{\Sigma}_{-1}(j),\bar{\Sigma}(j),Om\bar{e}ga(j),\bar{\omega}(j))$.

3.3 Calibration

To understand the macroeconomic implications of rationally inattentive multi-region firms, I calibrate the model using U.S. data as of 2012. The model is at a monthly frequency. I consider an economy with 12 regions[11], so $n = 12$. Each region represents one of the Federal Reserve districts and its boundaries and populations in 2012 are shown in Figure 1. I calibrate the regional population shares, $\{\zeta_l\}_{l\in\{1,...,12\}}$, using the BEA population data.

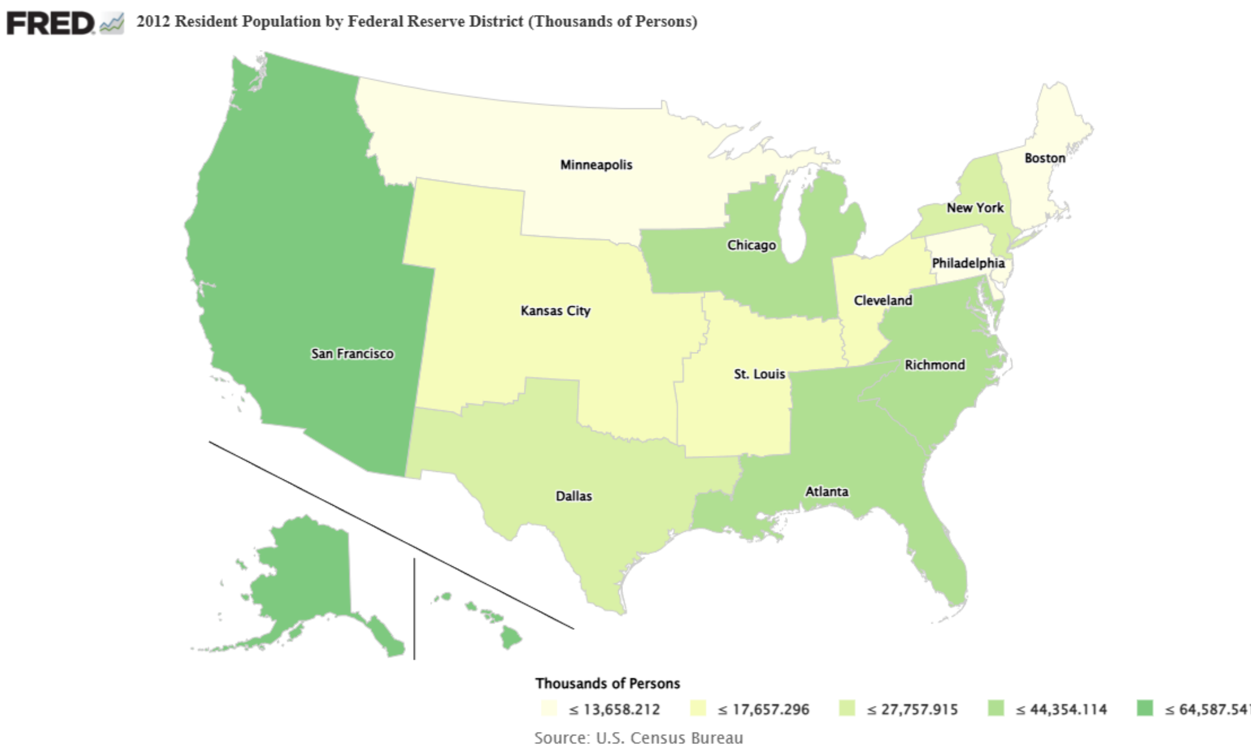


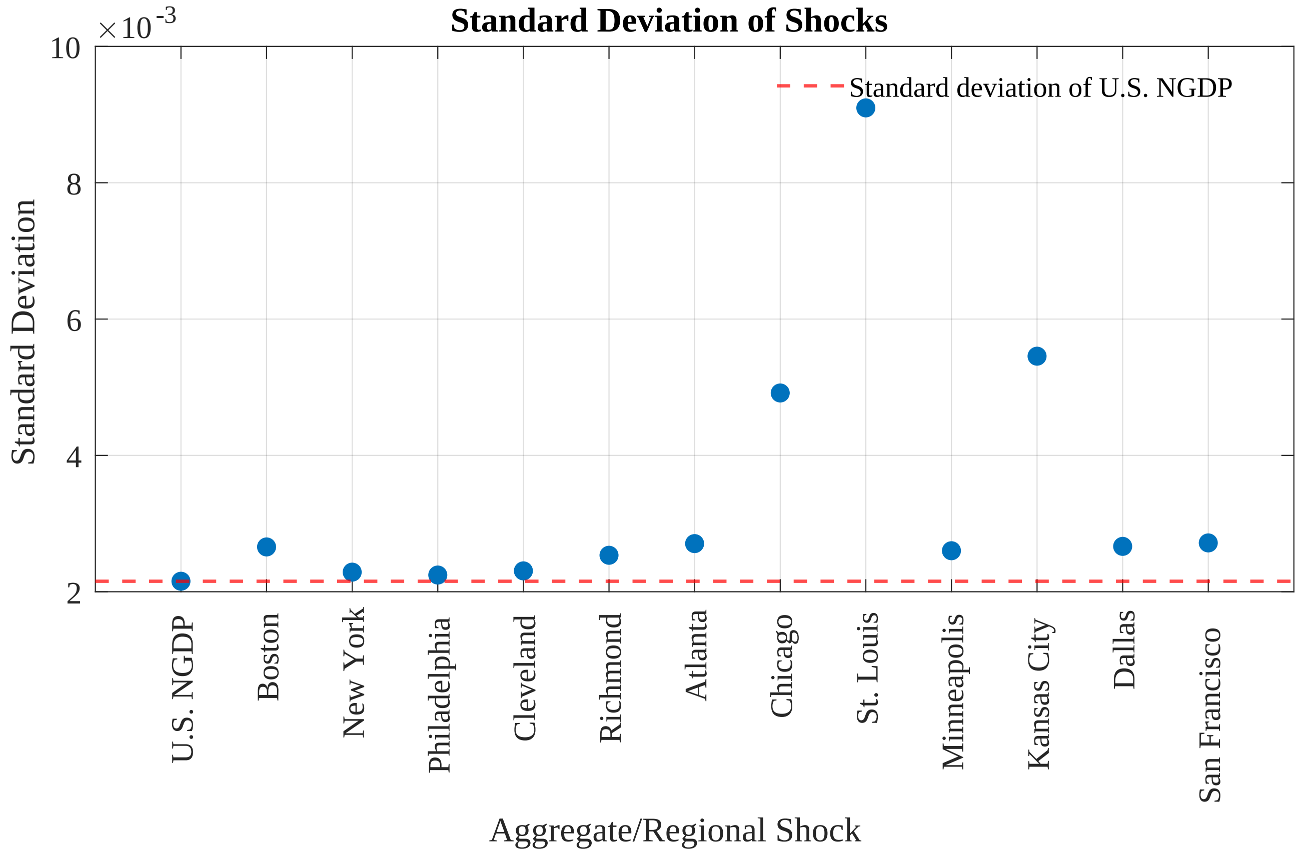
Figure 1: 2012 Resident Population by Federal Reserve districts (Thousands of Persons). Source: St. Louis FRED.

To calibrate the standard deviation of the aggregate shock, σ_u, I interpolate the U.S. nominal GDP using a spline routine to get nominal GDP monthly growth rates and use its standard deviation as σ_u. To calibrate the standard deviation of region specific markup shocks, $\{\sigma_l\}_{l\in\{1,...,12\}}$, I calculate regional employment levels using the QCEW, seasonally adjust them using X13, and then

[11] Ideally, one would like to have the most disaggregated geographic units as possible. However, computational issues start arising. First, recall that the number of types of firms is given by $2^n - 1$. Furthermore, with $n+1$ shocks, the prior variance-covariance matrix is a state variable, with $n(n+1)/2$ different variables, and the posterior variance-covariance matrix, which is a choice variable for the firm, also has $n(n+1)/2$ different variables. Hence, the number of potential problems to solve increase exponentially and the dynamic rational inattention problem becomes more complex.

use the regional standard deviation of the seasonally adjusted monthly employment growth rate as σ_l, $l \in \{1,\ldots,12\}$. Figure 2 shows the calibrated values for each one of these standard deviations. Appendices I.2 and I.3 provide a detailed description of their construction.

Figure 2: Calibrated values for standard deviation of shocks



Notes: Calibrated values of $\sigma_u = 0.00223$ and $\{\sigma_l\}_{l\in\{1,\ldots,12\}}$.

On the demand side of the economy, we need to calibrate the discount factor, β, the regional steady state markups, $\{\Lambda_l\}_{l=1,\ldots,12}$, and regional taste shifters for each retailer, $\{\theta_{lj}\}_{l\in\{1,\ldots,12\},j\in[0,1]}$. I set $\beta = (\frac{1}{1+0.02})^{1/12}$ which is how β is related to nominal interest rates in the steady state, assumed to be 2% annual. I set $\Lambda_l = {}^{4.5}/_{(4.5-1)}, \forall\, l \in \{1,\ldots,12\}$, where 4.5 is the elasticity of substitution across retailers within a region and taken from Hottman (2021), which estimates the elasticity of substitution across stores within a county using the U.S. NielsenIQ data.

Finally, I use the U.S. NielsenIQ scanner data to calibrate the share of each type of firm $\{\varphi_h\}_{h\mapsto\iota(h)}$, $\{\theta_{lj}\}_{l\in\{1,\ldots,12\},j\in[0,1]}$, and κ. The U.S. NielsenIQ scanner data contain weekly scanner prices and quantities for different products (UPCs) for retail stores in the United States. I work with a subset of product categories, products, and stores of the data for the period of 2006 to 2019. Broadly, I focus on food stores in the United States, product categories that compose the the BLS U.S. CPI food at home, products that are widely available in stores within each year, and retailers that are present in at least 10 out of the 14 years of data. A detailed description of the data and the data cleaning process are in Appendix I.1. Then, $\{\varphi_h\}_{h\mapsto\iota(h)}$ is calibrated as the number of firms that are type ι divided by the total number of firms used in the empirical analysis in 2012. $\{\theta_{lj}\}_{l\in\{1,\dots,12\},j\in[0,1]}$ are chosen to match the consumption expenditure share for each type of firm in each one of the regions, where I assume that $\int_0^1 \theta_{lj}\,dj = 1, \forall l \in [n]$. To go from consumption expenditure share of *types* of firms to individual firms j, I use the calibrated values of $\{\varphi_h\}_{h\mapsto\iota(h)}$. A detailed description of the calibration of $\{\theta_{lj}\}_{l\in\{1,\dots,12\},j\in[0,1]}$ is in Appendix I.4.

Parameter	Description	Value	Explanation
n	Number of regions	12	Federal Reserve Districts
$\{\zeta_l\}_{l\in[n]}$	Population in l		2012 Census estimate
σ_u	Std. of monetary shock		$\sigma(\Delta \log \text{NGDP}_t)$, 1990:01-2019:10
$\{\sigma_l\}_{l\in[n]}$	Std. of regional shocks		$\sigma(\Delta \log \text{Employment}_{lt})$, QCEW, 1990:01-2019:10
β	Discount factor	$(\frac{1}{1+0.02})^{1/12}$	$\beta = (\frac{1}{1+i})^{1/12}$
$\{\Lambda_l\}_{l\in[n]}$	Regional steady state markup	$\frac{4.5}{4.5-1}$	Hottman (2021)
$\{\varphi_h\}_{h\mapsto\iota(h)}$	Share of each type of firm		2012 U.S. NielsenIQ Data
$\{\theta_{lj}\}_{l\in[n],j\in\{0,1\}}$	Firms' regional taste shifters		2012 U.S. NielsenIQ Data

Table 1: Externally calibrated parameters.

The model has one internally calibrated parameter, the fixed capacity κ. I calibrate it using the following model moment to match its data counterpart:

$$m(\kappa) = \frac{1}{N_t + N_j} \sum_t \sum_j \left(\frac{var_{jt}(p_{ljt})}{var_t(p_{ljt})} \right)$$

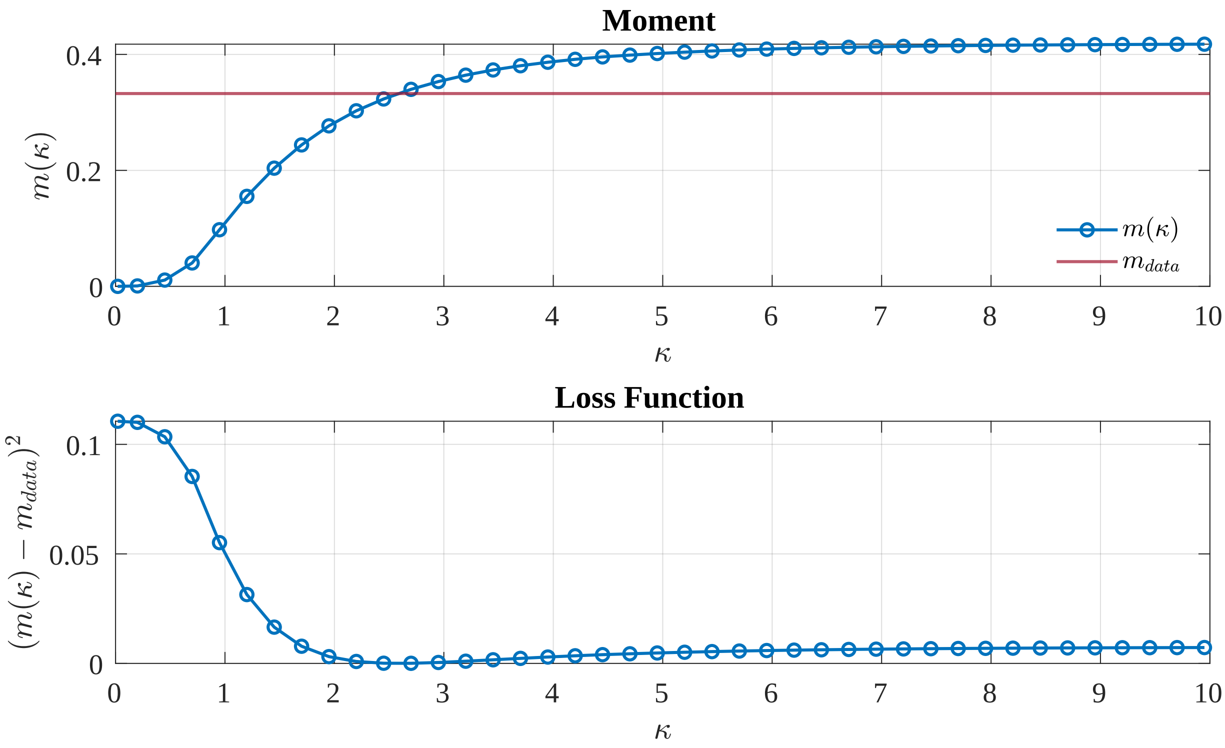
where j is a firm, l is a location, t is a time period. $var_{jt}(p_{jt})$ is the within-chain, within-period variance of prices across locations. That is, $var_{jt}(p_{ljt}) = \frac{1}{|\mathcal{L}_j|}\sum_l (p_{ljt} - \frac{1}{|\mathcal{L}_j|}\sum_l p_{ljt})^2$, where $|\mathcal{L}_j|$ is the number of regions where j operates. $var_t(p_{jt})$ is the within-period variance of prices across chains and locations. That is, $var_t(p_{ljt}) = \frac{1}{\sum_j |\mathcal{L}_j|}\sum_j\sum_l (p_{ljt} - \frac{1}{\sum_j |\mathcal{L}_j|}\sum_j\sum_l p_{ljt})^2$. I choose κ to solve the following minimization problem:

$$\min_{\kappa} g(\kappa) \equiv (m(\kappa) - m_{\text{data}})^2$$

where m_{data} is the data moment. A detailed description of how I calculate m_{data} is in Appendix I.5.

Figure 3 shows how the model moment and the loss function vary as I increase κ. The top figure shows that the moment increases monotonically with κ and it reflects the fact that regardless of firm type, as you increase κ the firm becomes less capacity constrained. As a result, its prices across regions become closer to their ideal flexible counterparts, increasing the between-region, within-chain price variance.

Figure 3: Model moment and loss function



Notes: Top: This figure plots the moment, $(m(\kappa))$, as you vary the capacity κ. It shows that $m(\kappa)$ is a strictly increasing function of κ for $\kappa \in [0.02, 10]$. Bottom: This figure plots the loss function $(m(\kappa) - m_{\text{data}})^2$ as you vary κ. Given that $m(\kappa)$ is strictly increasing in $[0.02, 10]$, there is a unique κ that minimizes $(m(\kappa) - m_{\text{data}})^2$, for $\kappa \in [0.02, 10]$. $m_{\text{data}} = 0.3325$.

3.4 Macroeconomic Implications

Having calibrated the model, I assess the quantitative implications of firms organization structure for decision-making for: (1) monetary non-neutrality; (2) propagation of regional markup shocks.

I compare the calibrated economy where firms decides at the headquarters level with an economy where multi-region firms decisions are made at the regional division level. In this case, the capacity is distributed according to each one of the region's sales share. That is, if a firm has capacity κ and total sales $\text{sales}_j = \sum_l \text{sales}_{lj}$, each region has $\kappa_{lj} = \frac{\text{sales}_{lj}}{\text{sales}_j}\kappa$. Under regional division decision-making, an expansionary monetary shock generates around six times larger monetary non-neutrality – as measured by the cumulative impulse response of aggregate GDP, than under headquarters decision-making. For the propagation of regional markup shocks, the quantitative implications depend on which region is shocked. Under regional division decision-making the aggregate GDP after a positive regional markup shock declines less compared to when decisions are made at the headquarters level. More importantly, under headquarters level decision-making, region-specific shocks spillover to other regions, a feature that is not present when decision-making occurs at the regional division level.

3.4.1 Monetary Non-Neutrality

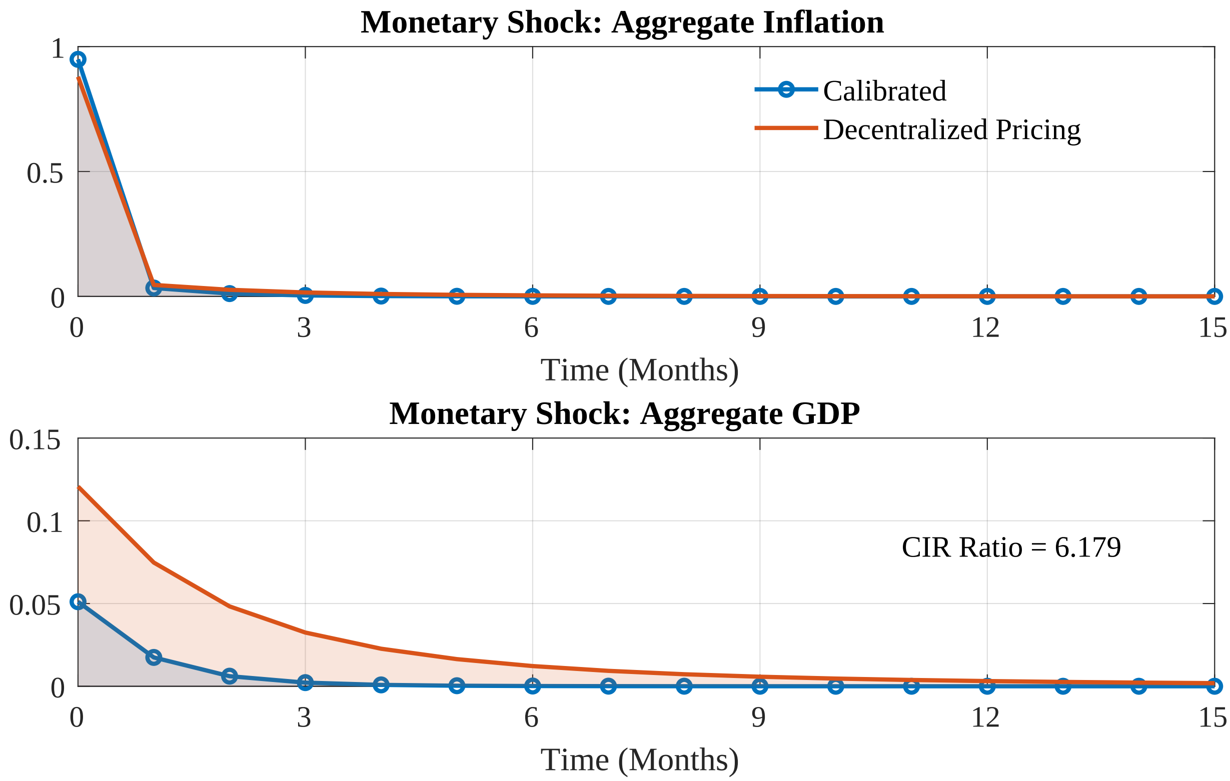
When decisions are made at headquarters, the real effects of monetary shocks on aggregate GDP are smaller than when decisions are made at the regional division level, as shown by the cumulative impulse response of aggregate GDP after an expansionary monetary shock.

Figure 4 illustrates that following a one standard deviation positive monetary shock, the cumulative impulse response of aggregate GDP under regional division decision-making (decentralized pricing) is six times larger than under headquarters decision-making (calibrated).

This result is linked to Corollary 1, where I show that under headquarters decision-making, firms' prices are more responsive to monetary shocks than under regional division decision-making. The introduction of persistent monetary shocks and positive discounting adds a dynamic component to the information acquisition process. Acquiring information today reduces future uncertainty, though the value is discounted over time. This force affects the *total* monetary non-neutrality under each one of the organization structures. However, for the *relative* monetary non-neutrality, the intuition is similar to the static case.

When decisions are made at headquarters, firms spread their attention across multiple signals. In contrast, under regional division decision-making, each division focuses on a single signal about its own demand but does so using only a fraction of the firm's total capacity. In the quantitative model, this leads to a dampened price response to aggregate shocks under regional division decision-making, thereby increasing monetary non-neutrality compared to headquarters decision-making.

Figure 4: Impulse response functions to a monetary policy shock



Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. Cumulative impulse responses (CIR) ratio denotes the ratio of cumulative impulse responses of the economy under regional division decision-making to the calibrated economy. The y-axis is in standard deviations of the shock.

3.4.2 Cross-Sectional Spillovers of Region-Specific Shocks

The calibrated economy amplifies the effects of regional markup shocks compared to an economy where firms' decision-making occurs at the regional division level. Importantly, when decisions are made at headquarters, regional markup shocks generate regional spillovers to other regions through firms' network of locations. In contrast, under region division level decision-making, there is no spillover.

Table 2 shows the ratio between cumulative impulse response of aggregate GDP under regional division decision-making relative to headquarters decision-making after a one standard deviation positive regional shock to the Fed district specified in each row. Regardless of the region that receives the shock, the contraction of aggregate GDP under regional division decision-making is lower than under headquarters decision-making. However, there's some heterogeneity in how large is the difference.

To illustrate how organizational structure affects the propagation of region-specific shocks, Figure 6 shows the spillover effects of a markup shock in the Atlanta Fed district across other districts under both headquarters and regional-division decision-making. First, this shock is contractionary for Atlanta Fed district under both headquarters and regional division decision-making, as expected. However, under headquarters decision-making, this shock *spills over* to other districts through firms' network of locations, as headquarters optimally choose noisy signals that don't allow them to perfectly identify the source of the shock. This spillover is quantitative relevant for some regions, as Cleveland and Dallas, whose GDP contractions are more than 10% of Atlanta's GDP contraction. This spillover is absent under regional division decision-making, as each regional division acquires a signle signal about its *own* demand conditions. Finally, we see regions (Boston and New York) for which there is no response, regardless of the organization structure. The reason for that is that there are no firms in the data that are present in Atlanta and in any of these regions.

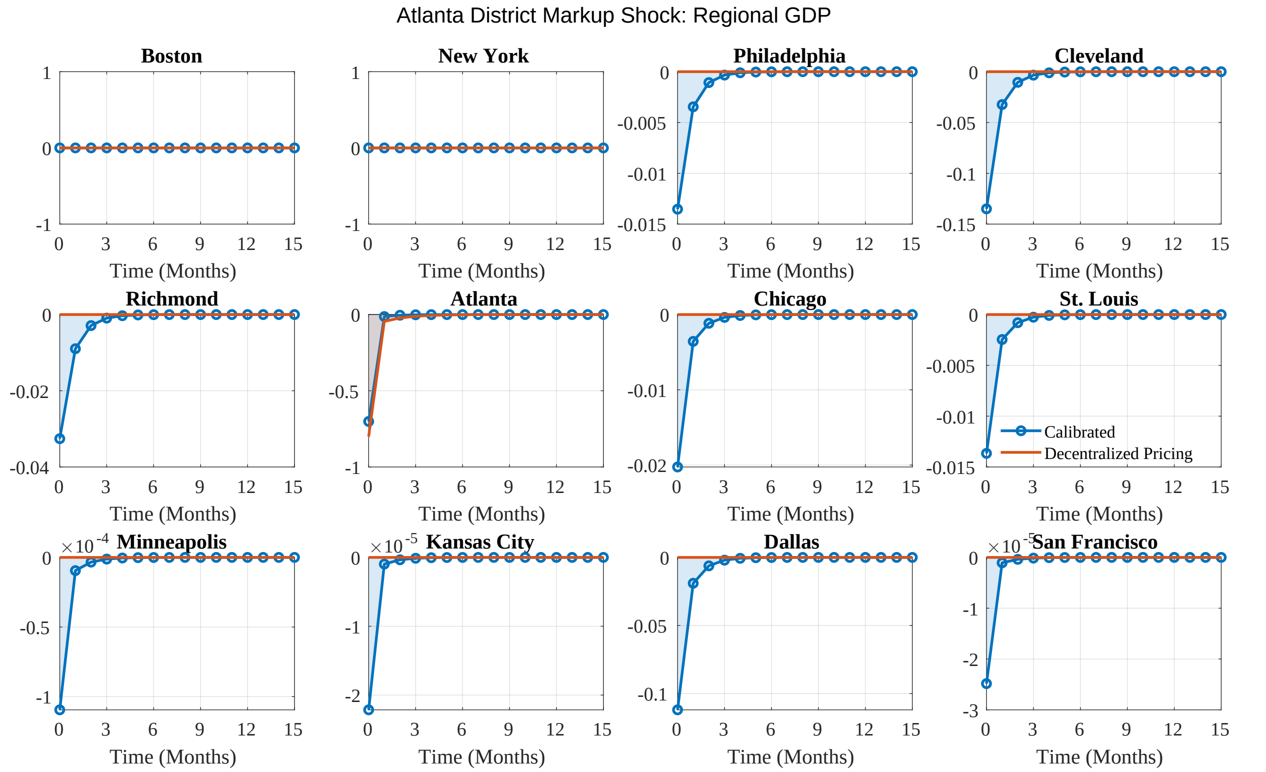
Table 2: Aggregate Effects of Regional Shocks

Fed District	CIR Ratio
Boston	0.9758
New York	0.9976
Philadelphia	0.9142
Cleveland	0.9691
Richmond	0.9274
Atlanta	0.9680
Chicago	0.9496
St. Louis	0.9015
Minneapolis	0.9722
Kansas City	0.9332
Dallas	0.9741
San Francisco	0.9952

Notes: This table shows the cumulative impulse response of aggregate GDP ratio between the economy under regional division decision-making and under headquarters decision-making. The decentralized pricing economy is an economy where each region of the multi-region firms solves its own problem. Mean: 0.956. Median: 0.968.

Figures J.1 to J.22 show the aggregate effects and the regional spillovers of region-specific shocks for shocks coming from each one of the other districts.

Figure 5: Impulse response functions to a markup shock in the Atlanta Fed district



Notes: This figure plots impulse response functions for regional inflation and regional GDP to a standard deviation positive monetary shock. The y-axis is in standard deviations of the shock.

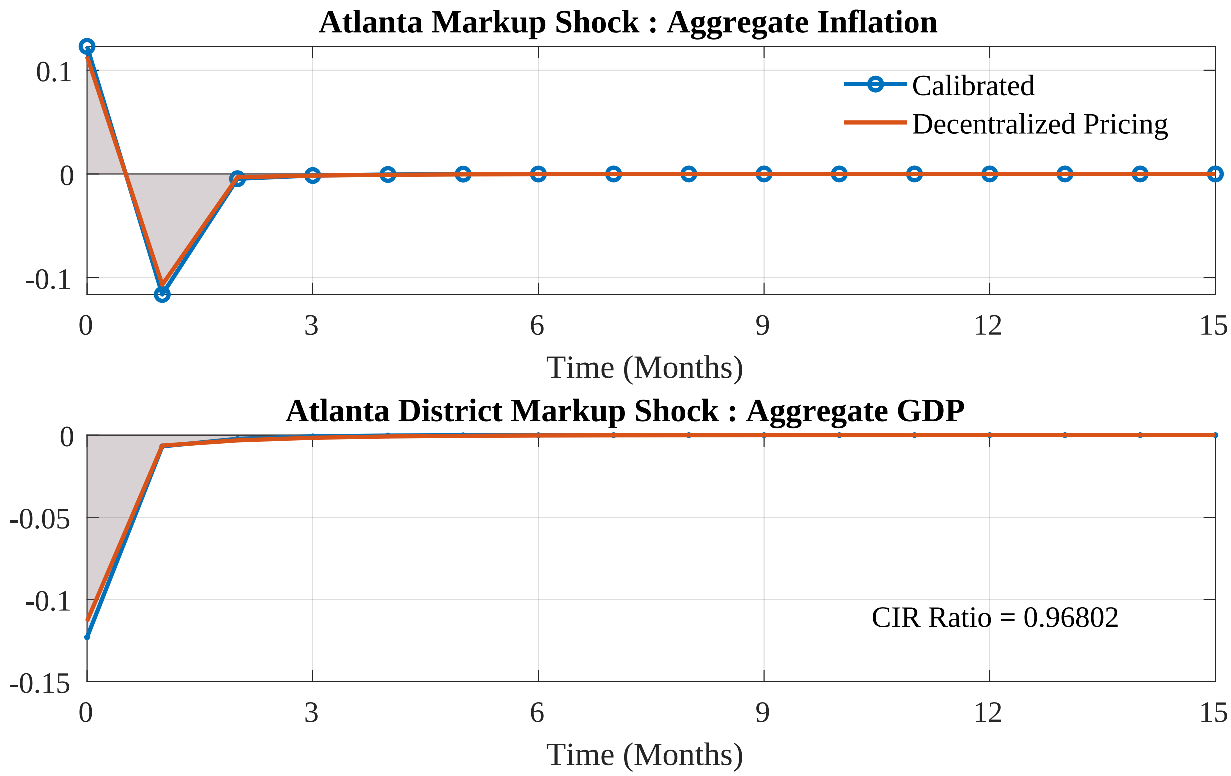
3.5 Model Validation

In this section, I present evidence supporting my theoretical framework. Building on Proposition 2, which shows that rational inattention reduces within-firm, between-region price dispersion compared to the full information case, I use the variation between regions and retailers in the NielsenIQ scanner data. Specifically, I show that for a given product, the relative price dispersion between regions is smaller when comparing prices within the same retail chain, as opposed to comparing prices across different chains. Then, I simulate data within my model and show that the model can qualitatively reproduce this result.

3.5.1 Between-County Relative Price Dispersion and Within-Chain Effect

Constructing the Sample. To compare prices of products between counties and assess how firms affect its dispersion, while accounting for the distance between counties, the relative prices dispersion must have both between-county and between-chain variation. I work with a subset of

Figure 6: Impulse response functions to a markup shock in the Atlanta Fed district



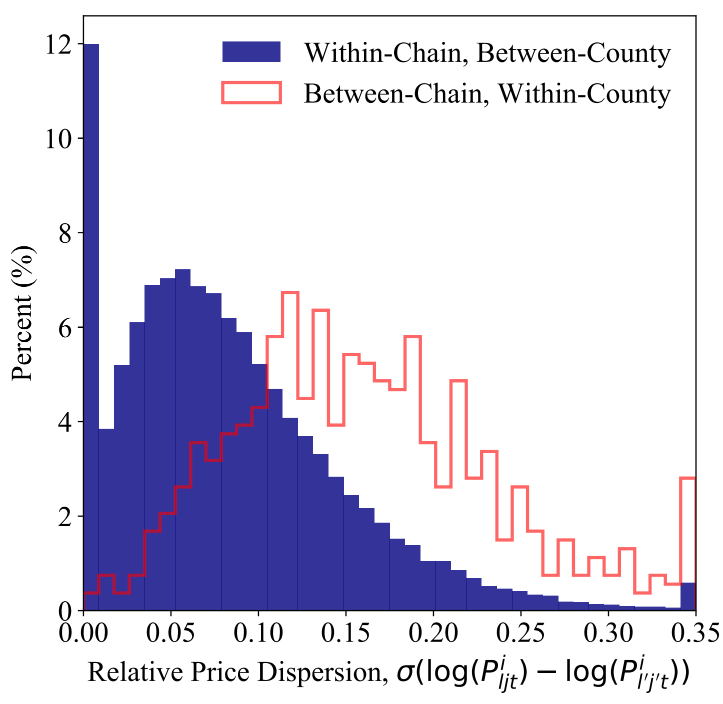
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of cumulative impulse response of aggregate GDP when decisions are made at the regional division relative to when decisions are made at headquarters. The y-axis is in standard deviations of the shock.

products that are widely available across stores within a year, and the prices I look at are related to the idea of *reference prices* (Eichenbaum, Jaimovich, and Rebelo, 2011). That is, the most often quoted price across all stores for a chain within a period of time. I consider a month. Finally, given the size of the data, and the fact that it is computationally infeasible to compare all possible pairs of prices across counties and chains, I work with a random sample of the data.

First, in a given year, I consider only UPCs sold in at least 70% of store-weeks. As a result, I consider goods that are widely available across stores, within a year. Then, given the subset of UPCs, I use the mode of prices for each UPC-Chain-County-Month. That is, I collapse data that are at the UPC-Store-Week level into data at UPC-Chain-County-Month level. This ensures that the variation I look at is at the Chain-County-Month level, disregarding variation that happens at the Store-Week level. I drop cases in which there is more than one mode. Finally, to compare prices of UPCs across counties and chains, for each UPC, I draw a random sample of 200 pairs of ordered pairs $((\text{county } l, \text{chain } j), (\text{county } l', \text{chain } j'))$. Then, for each pair, I calculate the relative price between $(\log P_{ljt} - \log P_{l'j't})$ for each period, and take its standard deviation over time, $\sigma(\log P_{ljt} - \log P_{l'j't})$. With this, I have both between-county, between-chain variation, and between-county, within-chain variation, while allowing to control for distance between counties. I use only pairs of prices such that I can calculate $(\Delta \log P_{ljt} - \Delta \log P_{l'j't})$ for at least 24 months.

Figure 7 displays the histogram of within-chain, between-county relative price dispersion, ie, $\sigma(\log P_{ljt} - \log P_{l'jt}), l \neq l'$, and the between-chain, within-county relative price dispersion, ie, $\sigma(\log P_{ljt} - \log P_{lj't}), j \neq j'$. As we can see, the within-chain, between-county relative price dispersion seems to be in general smaller than between-chain, within-county relative price dispersion. Figure J.24 plots the histogram considering only pairs of prices between counties that are at least 150 miles away for the within-firm, between-county relative price dispersion, displaying the same qualitative results.

Figure 7: Histogram of relative price dispersion



Notes: This figure plots the histogram of within-chain, between-county relative price dispersion and between-chain, within-county relative price dispersion.

Regression Specification. To uncover the effect of comparing prices between counties, but within the same retailer, while controlling for distance between the counties, I run the following cross-sectional gravity type regression pooled across all UPCs, which are indexed by i:

$$\sigma(\log P^i_{ljt} - \log P^i_{l'j't}) = \beta_0 \log D_{l,l'} + \beta_1 \mathbb{1}\{j = j'\} + \beta_2 \log D_{l,l'} \mathbf{1}\{j = j'\} + \beta_3 \mathbf{1}\{\text{Same Fed District}\} + \text{FE}_{l\times \text{m(i)}} + \text{FE}_{l'\times \text{m(i)}} + \epsilon^i_{((l,j),(l',j'))} \tag{32}$$

where $D_{l,l'}$ is the distance between counties l and l'[12], $\mathbf{1}\{j = j'\}$ is a dummy that takes the value of one, if $j = j'$ and zero otherwise, and $\mathbf{1}\{\text{Same Fed District}\}$ is a dummy that takes the value of one if l and l' belong to the same Federal Reserve District. $\text{FE}_{l\times \text{m(i)}}$ is a fixed effect for the county l and the module for good i. The coefficient of interest is β_1 and captures how comparing prices between counties for the same chain affects the relative price dispersion, after controlling for distance between counties. The coefficient β_0 captures how distance affects the relative price dispersion for a given good and it proxies shipping costs, differences in preferences that are function of distance, etc.

Results. Column (6) of Table 3 reports the results from estimating Equation (32) and shows that the relative price dispersion between counties for the same good decreases when we compare prices within chain, even after controlling for the distance between pairs of counties. It also shows that even when we compare prices within chain, relative price dispersion increases with the distance between counties. This evidence corroborates the findings from DellaVigna and Gentzkow (2019) that documents that chains display almost uniform prices across markets. At the same time, it is also consistent with the idea that between-county relative price dispersion increases with distance, as in Engel and Rogers (1996). This evidence is consistent with firm-level constraints that lower the variability of relative prices between regions.

[12] If $l = l'$, I assume $D_{l,l'} = 1$

Table 3: Between-County and Between-Chain Relative Price Dispersion

	(1)	(2)	(3)	(4)	(5)	(6)
log(distance)	0.011***			0.004***	0.003***	0.001***
	(0.000)			(0.000)	(0.000)	(0.000)
1(Same Chain)		-0.077***		-0.074***	-0.074***	-0.179***
		(0.002)		(0.002)	(0.002)	(0.004)
1(Same District)			-0.022***		-0.002***	-0.001**
			(0.001)		(0.000)	(0.000)
log(distance)1(Same Chain)						0.019***
						(0.001)
N. of obs.	724,893	724,893	724,893	724,893	724,893	724,893
R-squared	.013494	.0699358	.0096799	.071097	.0711378	.074855
Adjusted R-squared	.0134912	.0699332	.0096772	.0710932	.0711326	.0748486
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Chain FE	No	No	No	No	No	No

Notes: Dependent variable: $\sigma(\log(P^i_{ljt}/P^i_{l'j't}))$. Standard errors clustered at module. 1(Same Chain) = 1 if $j = j'$. 1(Same District) = 1 if l and l' belong to the same Federal Reserve district.

Since the results in Table 3 rely on a specific random sample of price pairs, I generate 50 random samples and rerun the regression from Column (5) of Equation (32) for each sample. Figures J.25 to J.27 illustrate the distribution of coefficients across these draws. Based on this analysis, I conclude that the point estimates in Table 3 are not driven by the specific random sample.

Model counterpart. With the calibrated model, I simulate a panel of firms' prices, draw a random sample of pairs of prices and then run the regression in column (5) of Table 3. While column (5) also controls for distance, in my model regions are completely segmented. Therefore, I do not need to control for distance, so I run the regression only against a dummy for same chain and a dummy for same district. Column (3) Table 4 shows that both same district and same chain dummies have negative effect on the relative price dispersion between regions, a result that is qualitatively consistent with Table 3.

Table 4: Model: Between-Region and Between-Chain Relative Price Dispersion

	(1)	(2)	(3)
1(Same Region)	-0.003***		-0.003***
	(0.000)		(0.000)
1(Same Chain)		-0.002***	-0.001***
		(0.000)	(0.000)
N. of obs.	1000000	1000000	1000000
R-squared	.5948288	.0013071	.5954165
Adjusted R-squared	.594819	.0012832	.5954064
Region FE	Yes	Yes	Yes

Notes: Dependent variable: $\sigma(\log(P^i_{ljt}/P^i_{l'j't}))$. Standard errors clustered at module. 1(Same Chain) = 1 if $j = j'$. 1(Same District) = 1 if l and l' belong to the same Federal Reserve district.

4 Conclusion

This paper develops a model to study how a firm's geographic dispersion affects and its organization structure affect information acquisition when this is costly. For firms operating in multiple regions under rational inattention, the level at which decisions are made–whether at the headquarters or regional divisions–plays a crucial role in shaping how firms form expectations about aggregate and region-specific shocks. The model provides a potential explanation for why firms may display low within-firm between-location relative price dispersion documented by the literature (DellaVigna and Gentzkow, 2019).

The framework can speak to recent trends highlighting the increased importance of national retail chains in the U.S. economy and how organization structure shapes a firm's price responses to different shocks. As firms expand nationally, those making decisions at the headquarters level will tend to focus more on overall demand conditions, resulting in stronger price response to monetary shocks. In contrast, those making decisions at the regional division will display weaker price responses to monetary shocks, as each one of the regional divisions will have lower capacity. Therefore, as firms become more national, it becomes increasingly important to understand whether decisions are made at the headquarters or regional division level, as the effects of monetary policy differ significantly under these two scenarios. An assumption for this conclusion is that the organization structure for decision-making does not vary with a firms geographic dispersion. A key direction for future research is to endogenize a firm's organization structure for information acquisition and pricing.

Furthermore, the proposed framework abstracts from other mechanisms that may drive endogenous propagation of region-specific shocks to focus specifically on how costly information acquisition alone can generate regional spillovers. It is important to assess the quantitative significance of the interaction between this new propagation mechanism and the usual ones for real business cycle comovement within a country like the United States. Among the simplifications, this paper assumes that monetary policy targets nominal aggregate GDP. Future research could explore the implications of this mechanism for welfare and optimal monetary policy.

References

ADAMS, B., AND K. R. WILLIAMS (2019): “Zone Pricing in Retail Oligopoly,” *American Economic Journal: Microeconomics*, 11(1), 12456.

AFROUZI, H. (2024): “Strategic Inattention, Inflation Dynamics, and the Nonneutrality of Money,” *Journal of Political Economy*, 132(10), 3378–3420.

AFROUZI, H., AND C. YANG (2021): “Dynamic Rational Inattention and the Phillips Curve,” Discussion paper.

BERAJA, M., E. HURST, AND J. OSPINA (2019): “The Aggregate Implications of Regional Business Cycles,” *Econometrica*, 87(6), 1789–1833.

BRODA, C., AND D. E. WEINSTEIN (2008): “Understanding International Price Differences Using Barcode Data,” Working Paper 14017, National Bureau of Economic Research.

DARUICH, D., AND J. KOZLOWSKI (2023): “Macroeconomic Implications of Uniform Pricing,” *American Economic Journal: Macroeconomics*, 15(3), 64–108.

DELLAVIGNA, S., AND M. GENTZKOW (2019): “Uniform Pricing in U.S. Retail Chains*,” *The Quarterly Journal of Economics*, 134(4), 2011–2084.

EICHENBAUM, M., N. JAIMOVICH, AND S. REBELO (2011): “Reference Prices, Costs, and Nominal Rigidities,” *American Economic Review*, 101(1), 23462.

ENGEL, C., AND J. H. ROGERS (1996): “How Wide Is the Border?,” *The American Economic Review*, 86(5), 1112–1125.

FULTON, C. (2022): “Choosing what to pay attention to,” *Theoretical Economics*, 17(1), 153–184.

GARCIA-LEMBERGMAN, E. (2020): “Multi-establishment Firms, Pricing and the Propagation of Local Shocks: Evidence from US Retail,” Working paper.

GIROUD, X., AND H. M. MUELLER (2019): “Firms’ Internal Networks and Local Economic Shocks,” *American Economic Review*, 109(10), 361749.

GUMPERT, A., H. STEIMER, AND M. ANTONI (2021): “Firm Organization with Multiple Establishments*,” *The Quarterly Journal of Economics*, 137(2), 1091–1138.

HAZELL, J., C. PATTERSON, H. SARSONS, AND B. TASKA (2023): "National Wage Setting," IZA Discussion Papers 16493, Institute of Labor Economics (IZA).

HOTTMAN, C. J. (2021): "Retail Markups, Misallocation, and Store Variety across U.S. Cities," Working paper.

HYUN, J., AND R. KIM (2019): "Spillovers through Multimarket Firms: The Uniform Product Replacement Channel," Available at SSRN: https://ssrn.com/abstract=3480961 or http://dx.doi.org/10.2139/ssrn.3480961.

KŐSZEGI, B., AND F. MATĚJKA (2020): "Choice Simplification: A Theory of Mental Budgeting and Naive Diversification*," *The Quarterly Journal of Economics*, 135(2), 1153–1207.

LEVY, D., M. BERGEN, S. DUTTA, AND R. VENABLE (1997): "The Magnitude of Menu Costs: Direct Evidence From Large U. S. Supermarket Chains," *The Quarterly Journal of Economics*, 112(3), 791–825.

LUCAS, R. E. (1972): "Expectations and the neutrality of money," *Journal of Economic Theory*, 4(2), 103–124.

MACKOWIAK, B., AND M. WIEDERHOLT (2009): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99(3), 769–803.

——— (2015): "Business Cycle Dynamics under Rational Inattention," *The Review of Economic Studies*, 82(4 (293)), 1502–1532.

MATĚJKA, F. (2015): "Rationally Inattentive Seller: Sales and Discrete Pricing," *The Review of Economic Studies*, 83(3), 1125–1155.

MIAO, J., J. WU, AND E. R. YOUNG (2022): "Multivariate Rational Inattention," *Econometrica*, 90(2), 907–945.

PASTEN, E., AND R. SCHOENLE (2016): "Rational inattention, multi-product firms and the neutrality of money," *Journal of Monetary Economics*, 80, 1–16.

SIMS, C. A. (2003): "Implications of rational inattention," *Journal of Monetary Economics*, 50(3), 665–690, Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information.

——— (2010): "Chapter 4 - Rational Inattention and Monetary Economics," vol. 3 of *Handbook of Monetary Economics*, pp. 155–181. Elsevier.

STEVENS, L. (2019): "Coarse Pricing Policies," *The Review of Economic Studies*, 87(1), 420–453.

YANG, C. (2022): "Rational inattention, menu costs, and multi-product firms: Micro evidence and aggregate implications," *Journal of Monetary Economics*, 128, 105–123.

Online Appendix

(Not for publication)

October 29, 2024

Table of Contents for the Online Appendix

A Additional Results and Extensions of the Static Model

In this section, I present additional results and extensions of the static model in Section 2.1 and the problem in Section 2.2.

A.1 Rational Inattention Problem When Decisions Are at the Headquarters Level

Proposition A.1. (Optimal Prices) Let $\{\underline{k}(\kappa), \{s_i\}_{i\in\{1,\dots,k\}}\}$ as in Proposition 1. Then, the optimal price for a firm j is given by:

1. When $k > \underline{k}(\kappa)$,

$$\vec{p}_j^* = \left(1 - e^{-2\kappa}\right)\left[(m + \frac{1}{k}\sum_{m=1}^{k}\lambda_k) + \sqrt{1+\frac{1}{k}}\nu_1\right]\mathbf{1}_{k\times 1} \tag{33}$$

   where $\nu_1 \sim \mathcal{N}(0, \frac{1}{e^{2\kappa}-1})$ and $\mathbf{1}_{k\times 1} = (1,\dots,1)'$. Therefore, the firm adopts a *uniform pricing* response policy.

2. When $k < \underline{k}(\kappa)$

$$\begin{aligned}
\vec{p}_j^* = &\left(1 - \frac{e^{-2\frac{\kappa}{k}}}{(1+k)^{1-\frac{1}{k}}}\right)\left[m + \frac{1}{k}\sum_{l\in\{1,\dots,k\}}\lambda_l + \sqrt{1+\frac{1}{k}}\nu_1\right]\mathbf{1}_{k\times 1} \\
&+ \frac{1}{2}\left(1 - \left[(k+1)\frac{1}{e^{2\kappa}}\right]^{\frac{1}{k}}\right)(\mathbf{e}_1 - \mathbf{e}_k)\left[(\lambda_1 - \lambda_k) + \sqrt{2}\nu_2\right] \\
&+ \sum_{i=3}^{k}\frac{i-1}{i}\left(1 - \left[(k+1)\frac{1}{e^{2\kappa}}\right]^{\frac{1}{k}}\right)\left(-\frac{1}{i-1}\sum_{m=1}^{i-2}\mathbf{e}_m + \mathbf{e}_{i-1} - \frac{1}{i-1}\mathbf{e}_k\right) \\
&\times\left[\lambda_{i-1} - \frac{1}{i-1}\sum_{m=2}^{i-1}\lambda_{m-1} - \frac{1}{i-1}\lambda_k + \sqrt{1+\frac{1}{i-1}}\nu_i\right]
\end{aligned} \tag{34}$$

   where $\nu_1 \sim \mathcal{N}\left(0, \frac{\frac{1}{k}\left[(k+1)\frac{1}{e^{2\kappa}}\right]^{\frac{1}{k}}}{(1+\frac{1}{k}) - \frac{1}{k}\left[(k+1)\frac{1}{e^{2\kappa}}\right]^{\frac{1}{k}}}\right)$, $\nu_i \sim \mathcal{N}\left(0, \frac{\frac{1}{k}\left[(k+1)\frac{1}{e^{2\kappa}}\right]^{\frac{1}{k}}}{\frac{1}{k} - \frac{1}{k}\left[(k+1)\frac{1}{e^{2\kappa}}\right]^{\frac{1}{k}}}\right)$, $2 \le i \le k$. Therefore, the firm adopts a *pricing-to-market pricing* response policy.

The price in a given location ℓ is given by $p_{\ell j}^* = \mathbf{e}_\ell' \vec{p}_j^*$.

Proposition A.2. (Capacity Increasing in k) Let the capacity function $\kappa(k) : \mathbb{R}_+ \to \mathbb{R}_+$ satisfy $\kappa'(k) > 0, \forall k > 0$, and let there exist $\bar{k} > 0$ such that $\frac{1}{2}\log(1+\bar{k}) = \kappa(\bar{k})$, with $\frac{1}{2}\log(1+k) < \kappa(k)$ for $k < \bar{k}$ and $\frac{1}{2}\log(1+k) > \kappa(k)$ for $k > \bar{k}$. Then, if $k \ge \bar{k}$, firm acquires only one signal, while if $k < \bar{k}$, the firm acquires k signals.

When the firm acquires only one signal, the signal is about the firm's overall demand:

$$s_1 = \mathbf{u}_1' \vec{x} + \nu_1, \quad \nu_1 \sim \mathcal{N}\left(0, \frac{1}{e^{2\kappa(k)} - 1}\right)$$

where the vector of loadings $\mathbf{u}_1 = (u_{11}, u_{12}, \ldots, u_{1,n+1})' \in \mathbb{R}^{n+1}$ is such that $||\mathbf{u}_1|| = 1$ and it satisfies

$$-u_{11} + \sum_{\ell=2}^{k+1} u_{1\ell} = 0 \tag{35}$$

$$u_{11} - k u_{1\ell} = 0, \quad 2 \leq \ell \leq k+1 \tag{36}$$

$$u_{1\ell} = 0, \quad k+1 < j \leq n+1 \tag{37}$$

When the firm acquires k signals, it acquires a signal about its overall demand:

$$s_1 = \mathbf{u}_1' \vec{x} + \nu_1, \quad \nu_1 \sim \mathcal{N}\left(0, \frac{\left(\frac{1}{k}\right)^{1-\frac{1}{k}} \left(\left(1 + \frac{1}{k}\right) e^{-2\kappa(k)}\right)^{\frac{1}{k}}}{\left(1 + \frac{1}{k}\right) - \left(\frac{1}{k}\right)^{1-\frac{1}{k}} \left(\left(1 + \frac{1}{k}\right) e^{-2\kappa(k)}\right)^{\frac{1}{k}}}\right)$$

with $\mathbf{u}_1 \in \mathbb{R}^{n+1}$ such that $||\mathbf{u}_1|| = 1$, satisfying Equations (35) to (37), and $k-1$ signals about relative demand with the following structure:

$$s_i = \mathbf{u}_i' \vec{x} + \nu_i, \quad \nu_i \sim \mathcal{N}\left(0, \frac{\left(\frac{1}{k}\right)^{1-\frac{1}{k}} \left(\left(1 + \frac{1}{k}\right) e^{-2\kappa(k)}\right)^{\frac{1}{k}}}{\left(\frac{1}{k}\right) - \left(\frac{1}{k}\right)^{1-\frac{1}{k}} \left(\left(1 + \frac{1}{k}\right) e^{-2\kappa(k)}\right)^{\frac{1}{k}}}\right), \quad 2 \leq i \leq k$$

and $\{\mathbf{u}_i\}_{i \in \{2,\ldots,k\}}, \mathbf{u}_i \in \mathbb{R}^{n+1}$ are such that $||\mathbf{u}_i|| = 1$, $\mathbf{u}_i \perp \mathbf{u}_{i'}, \forall i, i' \in \{1, \ldots, k\}, i \neq i'$, and they satisfy

$$\sum_{m=2}^{k+1} u_{i\ell} = 0 \tag{38}$$

$$u_{i1} = 0 \tag{39}$$

$$u_{i\ell} = 0, \; k+1 < j \leq n+1 \tag{40}$$

for $i \in \{2, \ldots, k\}$

A function that satisfies Proposition A.2 is $\kappa(k) = \frac{1}{2}\log(a + bk)$, with $a > 1$, $b \in (0,1)$.

Proposition A.3. (Optimal Information Structure With Different Prior Variance for Aggregate and Regional Shocks) Assume that the firm has a different prior variance regarding the aggregate shock and the regional shocks. That is,

$$\Sigma_{-1} = \begin{bmatrix} \sigma_m^2 & \mathbf{0}'_{1\times n} \\ \mathbf{0}_{n\times 1} & \sigma_\lambda^2 \mathbf{I}_{n\times n} \end{bmatrix}$$

Then the solution to the problem in Equation (5)-Equation (8) is such that there exists a threshold $\underline{k}(\kappa,\sigma_m^2,\sigma_\lambda^2) = (e^{2\kappa}-1)\frac{\sigma_\lambda^2}{\sigma_m^2}$ such that if $k > \underline{k}(\kappa,\sigma_m^2,\sigma_\lambda^2)$, the firm acquires *one* signal, while if $k \leq \underline{k}(\kappa,\sigma_m^2,\sigma_\lambda^2)$, the firm acquires k signals. When $k > \underline{k}(\kappa,\sigma_m^2,\sigma_\lambda^2)$, the signal that the firm acquires has the following structure:

$$s_1 = \mathbf{u}_1'\vec{x} + \nu_1, \quad \nu_1 \sim \mathcal{N}\left(0, \frac{1}{e^{2\kappa}-1}\right) \tag{41}$$

where the vector of loadings $\mathbf{u}_1 = (u_{11}, u_{12}, \ldots, u_{1,n+1})' \in \mathbb{R}^{n+1}$ is such that $||u_1|| = 1$ and it satisfies

$$-\sigma_\lambda u_{11} + \sum_{\ell=2}^{k+1} \sigma_m u_{1\ell} = 0 \tag{42}$$

$$-\sigma_\lambda u_{11} + k\sigma_m u_{1\ell} = 0,\ 2 \leq \ell \leq k+1 \tag{43}$$

$$u_{1\ell} = 0,\ k+1 < j \leq n+1 \tag{44}$$

When $k \leq \underline{k}(\kappa,\sigma_m^2,\sigma_\lambda^2)$, the signals that the firm acquire consist of one signal with the following structure:

$$s_1 = \mathbf{u}_1'\vec{x} + \nu_1, \quad \nu_1 \sim \mathcal{N}\left(0, \frac{\left(\frac{1}{k}\sigma_\lambda^2\right)^{\frac{k-1}{k}} \times \left[\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)e^{-2\kappa}\right]^{\frac{1}{k}}}{\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right) - \left(\frac{1}{k}\sigma_\lambda^2\right)^{\frac{k-1}{k}} \times \left[\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)e^{-2\kappa}\right]^{\frac{1}{k}}}\right) \tag{45}$$

with $\mathbf{u}_1 \in \mathbb{R}^{n+1}$ such that $||\mathbf{u}_1|| = 1$, satisfying Equations (42) to (44), and $k-1$ signals with the following structure:

$$s_i = \mathbf{u}_i'\vec{x} + \nu_i, \quad \nu_i \sim \mathcal{N}\left(0, \frac{\left(\frac{1}{k}\sigma_\lambda^2\right)^{\frac{k-1}{k}} \times \left[\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)e^{-2\kappa}\right]^{\frac{1}{k}}}{\left(\frac{1}{k}\sigma_\lambda^2\right) - \left(\frac{1}{k}\sigma_\lambda^2\right)^{\frac{k-1}{k}} \times \left[\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)e^{-2\kappa}\right]^{\frac{1}{k}}}\right),\ 2 \leq i \leq k \tag{46}$$

and $\{\mathbf{u}_i\}_{i\in\{2,\dots,k\}}, \mathbf{u}_i \in \mathbb{R}^{n+1}$ are such that $||\mathbf{u}_i|| = 1$, $\mathbf{u}_i \perp \mathbf{u}_{i'}, \forall i, i' \in \{1,\dots,k\}, i \neq i'$, and they satisfy

$$\sum_{m=2}^{k+1} u_{i\ell} = 0 \tag{47}$$

$$u_{i1} = 0 \tag{48}$$

$$u_{i\ell} = 0,\ k+1 < j \leq n+1 \tag{49}$$

for $i \in \{2,\dots,k\}$[13]

B Proofs

In this section, I show the proofs for the propositions and corollaries. I proceed by first showing in Appendix B.1 the results under headquarters decision-making, then the in Appendix B.2 results under regional division decision-making.

B.1 Headquarters Decision-Making

B.1.1 Proof of Proposition 1

Before solving the LQG-RI problem, let's first write the firm's objective function. Take a firm that is in k regions, $j \in [k] = \{1,2,\dots,k\}$. Note that:

$$\sum_{l\in\{1,\dots,k\}} \frac{B}{2}(p_{lj} - p^{\diamond}_{lj})^2 = (\vec{p}_j - \mathbf{H}'_j\vec{x})'\text{diag}(B/2)(\vec{p}_j - \mathbf{H}'_j\vec{x})$$

where diag($B/2$) is a diagonal matrix with diagonal values being $B/2$. Now, given a history of signals $S^0_j = S_{j0} \cup S^{-1}$, the optimal price under imperfect information

$$\vec{p}^*_j \equiv \text{argmax}_{\vec{p}_j} - \mathbb{E}\left[(\vec{p}_j - \vec{p}^{\diamond}_j)'\text{diag}(B/2)(\vec{p}_j - \vec{p}^{\diamond}_j)|S^0_j\right]$$
$$\vec{p}^*_j = \mathbb{E}\left[\vec{p}^{\diamond}_j|S^0_j\right]$$

[13] While equations (42)-(44) and (47)-(49) do not generate necessary an orthonormal basis, since the $\tilde{\tilde{\Omega}}$ is real and symmetric, the Spectral theorem guarantees the existence of an orthonormal basis. Then, one can use the Gram-Schmidt process to find an orthonormal basis.

Now, plugging $\vec{p}_j^*$ into $\mathbb{E}\big[(\vec{p}_j - \vec{p}_j^{\diamond})'\text{diag}(B/2)(\vec{p}_j - \vec{p}_j^{\diamond})|S_j^0\big]$ gives us

$$\mathbb{E}\big[(\mathbb{E}\big[\vec{p}_j^{\diamond}|S_j^0\big] - \vec{p}_j^{\diamond})'\text{diag}(B/2)(\mathbb{E}\big[\vec{p}_j^{\diamond}|S_j^0\big] - \vec{p}_j^{\diamond})|S_j^0\big]$$

Now, I perform a series of matrix operations to transform this into the objective function in Afrouzi and Yang (2021).

$$\begin{aligned}
\mathbb{E}\big[(\mathbb{E}\big[\vec{p}_j^{\diamond}|S_j^0\big] - \vec{p}_j^{\diamond})'\text{diag}(B/2)(\mathbb{E}\big[\vec{p}_j^{\diamond}|S_j^0\big] - \vec{p}_j^{\diamond})|S_j^0\big] &\overset{(1)}{=} \mathbb{E}\big[(\mathbb{E}\big[\mathbf{H}_j'\vec{x}|S_j^0\big] - \mathbf{H}_j'\vec{x})'\text{diag}(B/2)(\mathbb{E}\big[\mathbf{H}_j'\vec{x}|S_j^0\big] - \mathbf{H}_j'\vec{x})|S_j^0\big] \\
&\overset{(2)}{=} \mathbb{E}\big[(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})'\mathbf{H}_j\text{diag}(B/2)\mathbf{H}_j'(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})|S_j^0\big] \\
&\overset{(3)}{=} \mathbb{E}\big[(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})'\tilde{\tilde{\Omega}}_j(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})|S_j^0\big] \\
&\overset{(4)}{=} \mathbb{E}\big[tr\big((\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})'\tilde{\tilde{\Omega}}_j(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})\big)|S_j^0\big] \\
&\overset{(5)}{=} \mathbb{E}\big[tr\big(\tilde{\tilde{\Omega}}(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})'\big)|S_j^0\big] \\
&\overset{(6)}{=} tr\big(\tilde{\tilde{\Omega}}_j\mathbb{E}\big[(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})'|S_j^0\big]\big) \\
&\overset{(7)}{=} tr\big(\tilde{\tilde{\Omega}}_j\Sigma_{j0}\big)
\end{aligned}$$

where in (1) I use $\vec{p}_j^{\diamond} = \mathbf{H}_j'\vec{x}$, (2) I factor $\mathbb{H}_j$ out, and (3) define $\tilde{\tilde{\Omega}}_j \equiv \mathbf{H}_j\text{diag}(B/2)\mathbf{H}_j'$. In (4), I use the fact that the trace of a scalar is the scalar, (5) $tr(AB) = tr(BA)$, (6) $tr(\mathbb{E}[A]) = \mathbb{E}[tr(A)]$, taking $\tilde{\tilde{\Omega}}$ out of the expectation operation. Finally, in (7) $\Sigma_{j0} \equiv \mathbb{E}\big[(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})(\mathbb{E}\big[\vec{x}|S_j^0\big] - \vec{x})'|S_j^0\big]$. Then, we can use $tr\big(\tilde{\tilde{\Omega}}\Sigma_0\big) = tr\big(\Sigma_0\tilde{\tilde{\Omega}}\big)$. The LQG-RI problem is given by

$$\begin{aligned}
&\max_{\Sigma_0} -\frac{1}{2}tr(\Sigma_0\tilde{\tilde{\Omega}}) \\
\text{s.t. } &\frac{1}{2}\ln\left(\frac{|\Sigma_{-1}|}{|\Sigma_0|}\right) \le \kappa \\
&\Sigma_{-1} - \Sigma_0 \succeq 0 \\
&0 \prec \Sigma_{-1} \preceq \infty
\end{aligned}$$

or

$$\begin{aligned}
&\max_{\Sigma_0} -\frac{1}{2}\left\{tr(\Sigma_0\tilde{\tilde{\Omega}}) + \tilde{\omega}\ln\left(\frac{|\Sigma_{-1}|}{|\Sigma_0|}\right)\right\} + \tilde{\omega}_j\kappa \\
\text{s.t. } &\Sigma_{-1} - \Sigma_0 \succeq 0 \\
&0 \prec \Sigma_{-1} \preceq \infty
\end{aligned}$$

where $\tilde{\omega}_j$ is the Lagrange multiplier associated with the capacity constraint, and I omit j, and $\frac{1}{2}\ln\left(\frac{|\Sigma_{-1}|}{|\Sigma_0|}\right) = \kappa$, as we can always decrese profit losses by decreasing posterior uncertainty about the shocks. Note that this is basically a static version of the DRIP problem outlined in Afrouzi and Yang (2021), with the $\tilde{\omega}_j$ being the marginal cost of information acquisition and an additional constraint. Let's look at $\tilde{\tilde{\Omega}}_j \equiv \mathbf{H}_j \text{diag}(B/2)\mathbf{H}'_j$

$$\tilde{\tilde{\Omega}}_j \equiv \mathbf{H}_j \text{diag}(B/2)\mathbf{H}'_j$$

Since $j \in [k]$,

$$\mathbf{H}'_j = \begin{bmatrix} 1 & 1 & 0 & \dots & \dots & \dots & 0 \\ 1 & 0 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 1 & \dots & 0 \end{bmatrix}_{k\times(n+1)} = \begin{bmatrix} \mathbf{1}_{k\times 1} & \mathbf{I}_{(k\times k)} & \mathbf{0}_{k\times(n-k)} \end{bmatrix} \tag{50}$$

where $\mathbf{1}_{k\times 1}$ is a vector of ones, $\mathbf{I}_{k\times k}$ is the identity matrix, and $\mathbf{0}_{k\times(n-k)}$ is a matrix of zeros. So,

$$\tilde{\tilde{\Omega}}_j = \mathbf{H}_j \text{diag}\big(B\big)\mathbf{H}'_j \tag{51}$$

$$= \begin{bmatrix} \mathbf{1}'_{1\times k} \\ \mathbf{I}_{k\times k} \\ \mathbf{0}'_{(n-k)\times k} \end{bmatrix} \text{diag}\big(B\big)\begin{bmatrix} \mathbf{1}_{k\times 1} & \mathbf{I}_{k\times k} & \mathbf{0}_{k\times(n-k)} \end{bmatrix} \tag{52}$$

$$= \begin{bmatrix} kB & B\mathbf{1}' & \mathbf{0}_{1\times(n-k)} \\ B\mathbf{1} & \text{diag}\big(B\big)_{k\times k} & \mathbf{0}_{k\times(n-k)} \\ \mathbf{0}_{(n-k)\times 1} & \mathbf{0}_{(n-k)\times k} & \mathbf{0}_{(n-k)\times(n-k)} \end{bmatrix} \tag{53}$$

$$\tilde{\tilde{\Omega}}_j = kB \underbrace{\begin{bmatrix} 1 & \frac{1}{k}\mathbf{1}'_{1\times k} & \mathbf{0}_{1\times(n-k)} \\ \frac{1}{k}\mathbf{1}_{k\times 1} & \text{diag}\big(\frac{1}{k}\big)_{k\times k} & \mathbf{0}_{k\times(n-k)} \\ \mathbf{0}_{(n-k)\times 1} & \mathbf{0}_{(n-k)\times k} & \mathbf{0}_{(n-k)\times(n-k)} \end{bmatrix}}_{\equiv \Omega(k)} \tag{54}$$

We can solve the following problem

$$\begin{aligned} &\max_{\Sigma_0} -\frac{1}{2}\Big\{tr(\Sigma_0\Omega(k)) - \omega \ln\Big(\frac{|\Sigma_{-1}|}{|\Sigma_0|}\Big)\Big\} \\ \text{s.t. } &\Sigma_{-1} - \Sigma_0 \succeq 0 \\ &0 \prec \Sigma_{-1} \preceq \infty \end{aligned}$$

by factoring out kB and defining $\omega = \frac{1}{kB}\tilde{\omega}$, with $\frac{1}{2}\ln\Big(\frac{|\Sigma_{-1}|}{|\Sigma_0|}\Big) = \kappa$. This problem can be solved using Lagrangian methods, in particular we can use the results from Afrouzi and Yang (2021). Using Theorem 2.1 from Afrouzi and Yang (2021), we get

$$\Sigma_0 = \omega \Sigma_{-1}^{\frac{1}{2}}\Big[\text{Max}\big(\Sigma_{-1}^{\frac{1}{2}}\Omega\Sigma_{-1}^{\frac{1}{2}}, \omega\big)\Big]^{-1}\Sigma_{-1}^{\frac{1}{2}} \tag{55}$$

where Max is the operator such that for any symmetric matrix $\mathbf{X}$ with spectral decomposition $\mathbf{X} = \mathbf{UDU}'$, $\text{Max}(\mathbf{X},\omega) \equiv \mathbf{U}\max(\mathbf{D},\omega)\mathbf{U}'$, where $\max(\mathbf{D},\omega)$ operates on every element on the diagonal.

Since I assume that $\Sigma_{-1} = \mathbf{I}$, we have

$$
\begin{aligned}
\Sigma_0 &= \omega\left[\mathrm{Max}\left(\Omega,\omega\right)\right]^{-1} \\
&\overset{(1)}{=} \omega\left[\mathbf{U}\max\left(\mathbf{D},\omega\right)\mathbf{U}'\right]^{-1} \\
\Sigma_0 &\overset{(2)}{=} \omega\mathbf{U}\max\left(\mathbf{D},\omega\right)^{-1}\mathbf{U}^{-1}
\end{aligned}
\tag{56}
$$

where in (1) we perform a spectral decomposition of Ω, noting that Ω is real and symmetric. In (2), we basically use the inverse property, $(AB)^{-1} = B^{-1}A^{-1}$, and the fact that given that Ω is real and symmetric, we can find U orthogonal, such that $\mathbf{U}^{-1} = \mathbf{U}'$. Using the capacity constraint and the assumption that $\Sigma_{-1} = \mathbf{I}$, which implies in $|\Sigma_{-1}| = 1$, we have

$$
|\Sigma_0| = e^{-2\kappa}
$$

Taking the determinant of Equation (56), we get

$$
\begin{aligned}
|\Sigma_0| &= \omega^{n+1}|\max(\mathbf{D},\omega)^{-1}| \\
e^{-2\kappa} &= \omega^{n+1}|\max(\mathbf{D},\omega)^{-1}|
\end{aligned}
\tag{57}
$$

where $|\mathbf{U}| = 1$, as it is an orthogonal matrix, $|AB| = |A||B|$ and $|cA_{n\times n}| = c^n|A|$. Now, let's calculate the eigenvalues of Ω. We want to find d such that $det(\Omega(k) - d\mathbf{I}) = 0$. First, let's write $\Omega(k) - d\mathbf{I}$ as a block matrix

$$
(\Omega(k) - d\mathbf{I}) = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}
$$

where $\mathbf{A}_{11} = 1-d$, $\mathbf{A}_{12} = (\frac{1}{k}\mathbf{1}'_{1\times k}, \mathbf{0}_{1\times(n-k)})_{1\times n}$, $\mathbf{A}_{21} = \begin{pmatrix} \frac{1}{k}\mathbf{1}_{k\times 1} \\ \\ \mathbf{0}_{(n-k)\times 1} \end{pmatrix}_{n\times 1}$, and

$\mathbf{A}_{22} = \begin{bmatrix} \text{diag}(\frac{1}{k}-d)_{k\times k} & \mathbf{0}_{k\times(n-k)} \\ \\ \mathbf{0}_{(n-k)\times k} & -d\mathbf{I}_{(n-k)\times(n-k)} \end{bmatrix}_{n\times n}$. Since $\mathbf{A}_{11}$ is invertible, we have that

$$\begin{aligned} det(\Omega(k)-d\mathbf{I}) &= det(A_{11}) \times det(A_{22} - A_{21}A_{11}^{-1}A_{12}) \\ &= det(1-d) \times det(A_{22} + (-\frac{1}{1-d}A_{21})A_{12}) \end{aligned}$$

Note that A_{22} is a square diagonal matrix with non-zero main diagonal elements. Therefore, A_{22} is invertible. Furthermore, A'_{12} and $(-\frac{1}{1-d}A_{21})$ are column vectors. Then, by the matrix determinant lemma, we have

$$det(A_{22} + (-\frac{1}{1-d}A_{21})A_{12}) = (1 + A_{12}A_{22}^{-1}(-\frac{1}{1-d}A_{21})) \times det(A_{22})$$

Let's calculate these objects. First, since A_{22} is a diagonal matrix, $det(A_{22}) = \left(\frac{1}{k}-d\right)^k \times (-d)^{(n-k)}$ and

$$\begin{aligned} (1 + A_{12}A_{22}^{-1}(-\frac{1}{1-d}A_{21})) &= 1 + (\frac{1}{k}\mathbf{1}'_{1\times k}, \mathbf{0}_{1\times(n-k)}) \begin{bmatrix} \text{diag}(\frac{1}{k}-d)_{k\times k} & \mathbf{0}_{k\times(n-k)} \\ \\ \mathbf{0}_{(n-k)\times k} & -d\mathbf{I}_{(n-k)\times(n-k)} \end{bmatrix}^{-1} (-\frac{1}{1-d}\begin{pmatrix} \frac{1}{k}\mathbf{1}_{k\times 1} \\ \\ \mathbf{0}_{(n-k)\times 1} \end{pmatrix}) \\ &= 1 - \frac{1}{1-d}(\frac{1}{k}\mathbf{1}'_{1\times k}, \mathbf{0}_{1\times(n-k)}) \begin{bmatrix} \text{diag}(\frac{1}{k}-d)^{-1}_{k\times k} & \mathbf{0}_{k\times(n-k)} \\ \\ \mathbf{0}_{(n-k)\times k} & -\frac{1}{d}\mathbf{I}_{(n-k)\times(n-k)} \end{bmatrix} \begin{pmatrix} \frac{1}{k}\mathbf{1}_{k\times 1} \\ \\ \mathbf{0}_{(n-k)\times 1} \end{pmatrix} \\ &= 1 - \frac{1}{1-d}(\frac{1}{k}\mathbf{1}'_{1\times k}, \mathbf{0}_{1\times(n-k)}) \begin{pmatrix} \frac{\frac{1}{k}}{\frac{1}{k}-d}\mathbf{1}_{k\times 1} \\ \\ \mathbf{0}_{(n-k)\times 1} \end{pmatrix} \\ &= 1 - \frac{1}{1-d}[\frac{1}{k}(\frac{\frac{1}{k}}{\frac{1}{k}-d}) \times k] \\ &= 1 - \frac{1}{1-d}\frac{1}{k}\frac{1}{\frac{1}{k}-d} \end{aligned}$$

So

$$
\begin{aligned}
det(A_{22}+(-\frac{1}{1-d}A_{21})A_{12}) &= (1-\frac{1}{1-d}\frac{1}{k}\frac{1}{\frac{1}{k}-d})\times\left(\frac{1}{k}-d\right)^{k}\times(-d)^{(n-k)}\\
&= \frac{1}{1-d}\frac{1}{\frac{1}{k}-d}((1-d)(\frac{1}{k}-d)-\frac{1}{k})\times\left(\frac{1}{k}-d\right)^{k}\times(-d)^{(n-k)}\\
&= \frac{1}{1-d}((1-d)(\frac{1}{k}-d)-\frac{1}{k})\times\left(\frac{1}{k}-d\right)^{k-1}\times(-d)^{(n-k)}
\end{aligned}
$$

and

$$
\begin{aligned}
det(\Omega(k)-d\mathbf{I}) &= (1-d)\times\frac{1}{1-d}((1-d)(\frac{1}{k}-d)-\frac{1}{k})\times\left(\frac{1}{k}-d\right)^{k-1}\times(-d)^{(n-k)}\\
&= ((1-d)(\frac{1}{k}-d)-\frac{1}{k})\times\left(\frac{1}{k}-d\right)^{k-1}\times(-d)^{(n-k)}\\
&= (-d-\frac{d}{k}+d^2)\times\left(\frac{1}{k}-d\right)^{k-1}\times(-d)^{(n-k)}\\
det(\Omega(k)-d\mathbf{I}) &= \left(\left(1+\frac{1}{k}\right)-d\right)\times\left(\frac{1}{k}-d\right)^{k-1}\times(-d)^{(n-k+1)}
\end{aligned}
$$

To find the eigenvalues, we find d such that

$$
\left(\left(1+\frac{1}{k}\right)-d\right)\times\left(\frac{1}{k}-d\right)^{k-1}\times(-d)^{(n-k+1)}=0 \tag{58}
$$

From Equation (58), we can see that there is one eigen values with value $\left(1+{}^{1}/_{k}\right)$, $k-1$ eigenvalues with values ${}^{1}/_{k}$, and $(n-k+1)$ eigenvalues with zero value. Now, let's find the eigenvectors associated with those eigenvalues. First, when $d=1+\frac{1}{k}$, we have

$$
\begin{bmatrix}
1-\left(1+\frac{1}{k}\right) & \frac{1}{k}\mathbf{1}' & \mathbf{0}_{1\times(n-k)}\\
\frac{1}{k}\mathbf{1} & \operatorname{diag}\left(\frac{1}{k}-\left(1+\frac{1}{k}\right)\right)_{k\times k} & \mathbf{0}_{k\times(n-k)}\\
\mathbf{0}_{(n-k)\times 1} & \mathbf{0}_{(n-k)\times k} & -\left(1+\frac{1}{k}\right)\mathbf{I}_{(n-k)\times(n-k)}
\end{bmatrix}
\begin{pmatrix} v_1\\ v_2\\ \vdots\\ v_n\\ v_{n+1}\end{pmatrix}
=
\begin{pmatrix} 0\\ 0\\ \vdots\\ 0\\ 0\end{pmatrix}
$$

which implies in

$$-\frac{1}{k}v_1+\frac{1}{k}\sum_{j=2}^{k+1}v_j=0,\quad \frac{1}{k}v_1-v_j=0,\forall j\leq k+1,\quad -\left(1+\frac{1}{k}\right)v_j=0,j>k+1 \tag{59}$$

Let $v_1=1$, then $v_j={}^1\!/k, j\leq k+1$, $v_j=0, j>k+1$. So, an eigenvector associated with this eigen values is $(1,\underbrace{\frac{1}{k},\ldots,\frac{1}{k}}_{k\text{ times}},\underbrace{0,\ldots,0}_{n-k\text{ times}})'$. Note that this is not a unit vector. Therefore, we can define the unit eigenvector associated with $d_1=(1+\frac{1}{k})$ to be

$$\mathbf{u}_1=\frac{1}{\sqrt{1+\frac{1}{k}}}\times(1,\underbrace{\frac{1}{k},\ldots,\frac{1}{k}}_{k\text{ times}},\underbrace{0,\ldots,0}_{n-k\text{ times}})' \tag{60}$$

When $d=\frac{1}{k}$, we have

$$\begin{bmatrix} 1-\frac{1}{k} & \frac{1}{k}\mathbf{1}' & \mathbf{0}_{1\times(n-k)} \\ \frac{1}{k}\mathbf{1} & \mathbf{0} & \mathbf{0}_{k\times(n-k)} \\ \mathbf{0}_{(n-k)\times 1} & \mathbf{0}_{(n-k)\times k} & -\left(\frac{1}{k}\right)\mathbf{I}_{(n-k)\times(n-k)} \end{bmatrix}\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \\ v_{n+1}\end{pmatrix}=\begin{pmatrix}0\\0\\\vdots\\0\\0\end{pmatrix}$$

which implies in

$$\left(1-\frac{1}{k}\right)v_1+\frac{1}{k}\sum_{j=2}^{k+1}v_j=0,\quad \frac{1}{k}v_1=0,\quad -\frac{1}{k}v_j=0,j>k+1$$

Hence, the eigenspace associated with the eigenvalue $d_2=\frac{1}{k}$ is spanned by $\{\mathbf{e}_i^{(n+1)}-\mathbf{e}_{k+1}^{(n+1)}\}_{i\in\{2,3,\ldots,k\}}$, where $\mathbf{e}_i^{(n+1)}$ is the standard basis vector of dimension $n+1$, with one in the i-th entry. These are $k-1$ eigenvectors. Note that $\{\mathbf{e}_i^{(n+1)}-\mathbf{e}_{k+1}^{(n+1)}\}_{i\in\{2,3,\ldots,k\}}$ is not an orthonormal basis. To construct an orthonormal basis, we can perform a Gram-Schmidt process to $\{\mathbf{e}_i^{(n+1)}-\mathbf{e}_{k+1}^{(n+1)}\}_{i\in\{2,3,\ldots,k\}}$, which will give us an orthonormal basis $\{\mathbf{u}_i\}_{i=2,\ldots,k}$. More specifically, the Gram-Schmidt process to $\{\mathbf{e}_i^{(n+1)}-$

$\mathbf{e}_{k+1}^{(n+1)}\}_{i\in\{2,3,\ldots,k\}}$ yields the following orthonormal basis $\{\mathbf{u}_i\}_{2\le i\le k}$:

$$\mathbf{u}_i=\frac{1}{\sqrt{1+\frac{1}{i-1}}}(e_i^{(n+1)}-\frac{1}{i-1}\sum_{m=2}^{i-1}e_m^{(n+1)}-\frac{1}{i-1}e_{k+1}^{(n+1)}),\, i\ge 2 \tag{61}$$

where $\sum_{m=2}^{i-1}e_m^{(n+1)}=0$ for $i=2$.

For $d=0$, we do not need to calculate the eigenvectors as firms never acquire information along the dimension spanned by the eigenvectors associated with the eigenvalue $d=0$.

Therefore,

$$\mathbf{D}=\mathrm{diag}(d_{ii})_{i\in\{1,\ldots,n+1\}},\quad d_{ii}=\begin{cases}1+\frac{1}{k} & ,i=1\\ \frac{1}{k} & ,2\le i\le k+1\\ 0 & ,k+1<i\le n+1\end{cases} \tag{62}$$

From Equation (57), and Equation (62), we have two cases: (1) $\omega\in(\frac{1}{k},1+\frac{1}{k}]$, and (2) $\omega\in(0,\frac{1}{k}]$. Note that $\omega>1+\frac{1}{k}$ implies that the firm won't acquire information at all, which is not optimal. Let's first consider the first case

Case 1: $\omega\in(\frac{1}{k},1+\frac{1}{k}]$**.** In this case, firm acquires one signal, as the marginal cost of information acquisition ω is higher than the marginal benefit of pay attention to signals with eigenvalue $1/k$. Also, when $\omega\in(\frac{1}{k},1+\frac{1}{k}]$, using Equation (57), we have

$$\omega=\left(1+\frac{1}{k}\right)e^{-2\kappa} \tag{63}$$

Finally, we can find (κ,k) such that $\omega\in(\frac{1}{k},1+\frac{1}{k}]$:

$$\omega\in(\frac{1}{k},1+\frac{1}{k}]\iff\omega>\frac{1}{k}\text{ and }\omega\le 1+\frac{1}{k}$$

$$\left(1+\frac{1}{k}\right)e^{-2\kappa}>\frac{1}{k}\iff k>e^{2\kappa}-1$$

and

$$\left(1+\frac{1}{k}\right)e^{-2\kappa}\le 1+\frac{1}{k}\iff e^{2\kappa}\ge 1$$

which is true for all $\kappa > 0$. Therefore, we have that a firm that operates in k regions pays attention to a single signal when $k > e^{2\kappa} - 1$.

Case 2: $\omega \in (0, \frac{1}{k}]$**.** In this case, the firm acquires k signals, as the marginal cost of information acquisition ω is smaller than the marginal benefit of paying attention to signals with eigenvalue of at least $1/k$. Also, when $\omega \in (0, \frac{1}{k}]$, using Equation (57), we have

$$\omega = \frac{1}{k}\left[(k+1)e^{-2\kappa}\right]^{\frac{1}{k}} \tag{64}$$

Finally, we can find (κ, k) such that $\omega \in (0, \frac{1}{k}]$:

$$\frac{1}{k}\left[(k+1)\frac{1}{e^{2\kappa}}\right]^{\frac{1}{k}} \leq \frac{1}{k} \iff k \leq e^{2\kappa} - 1$$

Therefore, with with cases 1 and 2, we have that a firm the operates in k regions that has κ capacity acquires one signal when $k > e^{2\kappa} - 1$, and acquires k signals when $k \leq e^{2\kappa} - 1$. By defining $\underline{k}(\kappa) = e^{2\kappa} - 1$, we have the first part of Proposition 1.

From Theorem 2.2 from Afrouzi and Yang (2021), let $\{d_i(k)\}_{1\leq i\leq n+1}$ be the set of eigenvalues of $\Omega(k)$ indexed in descending order. Let $\{\mathbf{u}_i(k)\}_{1\leq i\leq n+1}$ be orthonormal eigenvectors that correspond to those eigenvalues. Then, firm j's posterior belief is spanned by the following $0 \leq k^+ \leq k$ signals

$$s_{ij}(k) = \mathbf{g}_i'(k)\vec{x} + \nu_{ij}(k), \quad 1 \leq i \leq k^+$$

where k^+ is the number of the eigenvalues that are at least as large as ω, and for $i \leq k^+$, $\mathbf{g}_i(k) \equiv \mathbf{u}_i(k)$ is the loading of signal i on $\vec{x}$, and $\nu_{ij} \sim \mathcal{N}(0, \frac{\omega}{d_i(k)-\omega})$ is the firm's rational inattention error in signal i that is orthogonal to $\vec{x}$ and all other rational inattention errors. Let's consider the two cases above again:

Case 1: $\omega \in (\frac{1}{k}, 1 + \frac{1}{k}]$**.** When $k > e^{2\kappa} - 1$, the firm acquires one signal. The eigenvector associated with $d_1(k)$ is $\mathbf{u}_1(k) = \frac{1}{\sqrt{1+\frac{1}{k}}}(1, \underbrace{\frac{1}{k}, \ldots, \frac{1}{k}}_{k-1 \text{ times}}, \underbrace{0, \ldots, 0}_{n-k+1 \text{ times}})'$. Therefore, the signal that the firm acquires has the following structure:

$$s_{1j}^{\text{HQ}}(k) = \frac{1}{\sqrt{1+\frac{1}{k}}}(1, \frac{1}{k}, \ldots, \frac{1}{k}, 0, \ldots, 0)\vec{x} + \nu_{ij}(k), \quad \nu_{1j} \sim \mathcal{N}\big(0, \frac{1}{e^{2\kappa} - 1}\big)$$

To provide a better idea about the interpretation of this signal, we can rewrite it in terms of firm's ideal prices:

$$s_{1j}^{\mathrm{HQ}}(k) = \frac{1}{\sqrt{1+\frac{1}{k}}}\Big[\frac{1}{k}\sum_{l\in[k]} p_l^{\diamond}\Big] + \nu_{ij}(k), \quad \nu_{1j} \sim \mathcal{N}\big(0, \frac{1}{e^{2\kappa}-1}\big)$$

We can see that $s_{1j}^{\mathrm{HQ}}(k)$ is a signal about firm's *average prices* across regions where it operate. Given the log-linear relationship between regional demand and prices, this can interpreted as a signal about the firm's *overall demand* across those regions.

Case 2: $\omega \in (0, \frac{1}{k}]$**.** When $k \leq e^{2\kappa} - 1$, the firm acquires k signals. The eigenvector associated with $d_1(k)$ is $\mathbf{u}_1(k) = \frac{1}{\sqrt{1+\frac{1}{k}}}(1, \underbrace{\frac{1}{k}, \ldots, \frac{1}{k}}_{k-1 \text{ times}}, \underbrace{0, \ldots, 0}_{n-k+1 \text{ times}})'$. The eigenvectors associated with the repeated eigenvalue $d_2 = \frac{1}{k}$ are given by $\{\mathbf{u}_i : \mathbf{u}_i = \frac{1}{\sqrt{1+\frac{1}{i-1}}}(e_i^{(n+1)} - \frac{1}{i-1}\sum_{m=2}^{i-1} e_m^{(n+1)} - \frac{1}{i-1}e_{k+1}^{(n+1)}), i \in \{2, \ldots, k\}\}$. Therefore, the signals that the firm acquires have the following structure:

1. One signal about its *overall demand*:

$$s_{1j}^{\mathrm{HQ}}(k) = \frac{1}{\sqrt{1+\frac{1}{k}}}(1, \frac{1}{k}, \ldots, \frac{1}{k}, 0, \ldots, 0)\vec{x} + \nu_{ij}(k), \quad \nu_{1j} \sim \mathcal{N}\big(0, \frac{1}{(k+1)^{1-\frac{1}{k}}(e^{2\kappa})^{\frac{1}{k}} - 1}\big)$$

2. $k-1$ signals about its regional *relative demands*:

$$s_{ij}^{\mathrm{HQ}}(k) = \frac{1}{\sqrt{1+\frac{1}{i-1}}}(p_{i-1}^{\diamond} - \frac{1}{i-1}\sum_{m=2}^{i-1} p_{m-1}^{\diamond} - \frac{1}{i-1}p_k^{\diamond}) + \nu_{ij}(k),$$

$$\nu_{ij}(k) \sim \mathcal{N}\left(0, \frac{1}{(k+1)^{-\frac{1}{k}}(e^{2\kappa})^{\frac{1}{k}} - 1}\right) \quad i \in \{2, \ldots, k\}$$

where I wrote both signal structure in terms of ideal prices. Recall that the $k-1$ are signals about *relative demands* because its eigenvectors are basically an orthonormal basis that span the same eigenspace as $\{p_i^{\diamond} - p_k^{\diamond}\}_{i\in\{1,\ldots,k-1\}}$, which are vectors of relative prices.

With cases 1 and 2, we have the second part of the Proposition 1. □

B.1.2 Proof of Proposition 2

This result follows from optimal prices in Proposition A.1 by taking the difference between p_{1j}^{HQ} and $p_{\ell j}^{\mathrm{HQ}}, \ell \in [k]$ and then calculating its variance, $var(p_{1j}^{\mathrm{HQ}} - p_{\ell j}^{\mathrm{HQ}})$ and comparing with $var(p_{1j}^{\diamond} - p_{\ell j}^{\diamond}) =$

$2, \ell \in [k]$. For a firm j in k regions, after some algebra, we get

$$var(p_{1j}^* - p_{\ell j}^*) = \begin{cases} 0 & , k > e^{2\kappa} - 1 \\ 2(1 - ((k+1)e^{-2\kappa})^{1/k}) & , k \leq e^{2\kappa} - 1 \end{cases}$$

For $k > e^{2\kappa} - 1$, it is easy to see that $var(p_{1j}^* - p_{\ell j}^*) = 0 < 2$. For $k \leq e^{2\kappa} - 1$,

$$2(1 - ((k+1)e^{-2\kappa})^{1/k}) \leq 2 \iff 1 - ((k+1)e^{-2\kappa})^{1/k} \leq 1 \iff ((k+1)e^{-2\kappa})^{1/k} \geq 0$$

which is always true for $k \geq 1, \kappa > 0$. Finally, with a slight abuse of notation, when $k \leq e^{2\kappa} - 1$,

$$\frac{\partial var(p_{1j}^* - p_{\ell j}^*)}{\partial k} = -\frac{2}{k}\Big[(k+1)e^{-2\kappa}\Big]^{\frac{1}{k}} \Big\{ \big(\frac{1}{k+1}\big) - \frac{1}{k}\big(\ln(k+1) - 2\kappa\big) \Big\} < 0, \forall k \geq 2$$

as $\big(\frac{1}{k+1}\big) - \frac{1}{k}\big(\ln(k+1) - 2\kappa\big) \geq 0$. To see this, note that

$$k \leq e^{2\kappa} - 1 \iff \ln(k+1) - 2\kappa < 0$$

and

$$\big(\frac{1}{k+1}\big) - \frac{1}{k}\big(\ln(k+1) - 2\kappa\big) > 0 \iff \underbrace{\big(\frac{k}{k+1}\big)}_{>0} > \underbrace{\big(\ln(k+1) - 2\kappa\big)}_{<0,\text{ as } k \leq e^{2\kappa}-1}$$

and within-firm relative price dispersion between region one and ℓ, $\ell \in [k]$ is decreasing in k. □

B.1.3 Proof of Proposition 3

This result follows from Proposition A.1 by taking derivatives of p_{lj}^{HQ} with relation to m. This results in

$$\frac{\partial p_{lj}^{\text{HQ}}}{\partial m} = \begin{cases} (1 - e^{-2\kappa}) & , k > \underline{k}(\kappa) \\ 1 - \frac{1}{(k+1)^{1-\frac{1}{k}}(e^{2\kappa})^{\frac{1}{k}}} & , k \leq \underline{k}(\kappa) \end{cases}$$

Then, for $k \leq \underline{k}(\kappa)$, we can calculate $\frac{\partial(\partial p_{lj}^{\text{HQ}}/\partial m)}{\partial k}$. First, let $f(k) \equiv (k+1)^{-(1-\frac{1}{k})}$. Then,

$$\frac{\partial(\frac{\partial p_{lj}^{\text{HQ}}}{\partial m})}{\partial k} = -\Big[f'(k)(e^{-\frac{2\kappa}{k}}) + f(k)(e^{-\frac{2\kappa}{k}})\Big(\frac{2\kappa}{k^2}\Big)\Big]$$

where $f'(k) = -f(k)\Big[\frac{1}{k^2}\ln(k+1) + (1-\frac{1}{k})\frac{1}{k+1}\Big]$. So,

$$\frac{\partial(\frac{\partial p_{lj}^{\text{HQ}}}{\partial m})}{\partial k} = -f(k)(e^{-\frac{2\kappa}{k}})\frac{1}{k}\Big[-\frac{1}{k}\ln(k+1) - \frac{(k-1)}{k+1} + \Big(\frac{2\kappa}{k}\Big)\Big]$$

Now, note that

$$\frac{\partial(\frac{\partial p_{lj}^{\text{HQ}}}{\partial m})}{\partial k} \geq 0 \iff -\frac{1}{k}\ln(k+1) - \frac{(k-1)}{k+1} + \Big(\frac{2\kappa}{k}\Big) \leq 0$$
$$\iff 2\kappa \leq \ln(k+1) + \frac{k(k-1)}{k+1}$$

Let $h(k) \equiv \ln(k+1) + \frac{k(k-1)}{k+1}$. Note that $h'(k) > 0, \forall k \geq 1$. For $\kappa \leq \frac{\ln(2)}{2}$, we have that $h(k) \geq 2\kappa, \forall k \geq 1$. Now, for $\kappa > \frac{\ln 2}{2}$ we have that $h(1) < 2\kappa$. However, since $h'(k) > 0 \forall k \geq 1$, $\exists \underline{k}^*$ such that $h(\underline{k}^*) = 2\kappa$, for which $k \leq \underline{k}^*$ implies in $h(k) \leq 2\kappa$, and for which $k > \underline{k}^*$ implies in $h(k) > 2\kappa$. Since $\frac{\partial(\frac{\partial p_{lj}^{\text{HQ}}}{\partial m})}{\partial k} \geq 0 \iff 2\kappa \leq \ln(k+1) + \frac{k(k-1)}{k+1}$, we have just proved the result. $\square$

B.1.4 Proof of Proposition 4

This result follows from Proposition A.1 by taking derivatives of p_{lj}^{HQ} with relation to λ_ℓ, $\ell \in [k]$. $\square$

B.2 Regional Division Decision-Making

B.2.1 Proof of Proposition 5

Before solving the LQG-RI problem, let's first write the firm's objective function. Take a regional division operating in l. First, note that $p_{lj}^{\diamond} = (\mathbf{e}_1^{(n+1)} + \mathbf{e}_{l+1}^{(n+1)})'\vec{x}$. Using similar arguments as in the Proof of Proposition 1, we have

$$\mathbb{E}\big[(\mathbb{E}[p_{lj}^{\diamond}|S_j^0] - p_{lj}^{\diamond})'B(\mathbb{E}[p_{lj}^{\diamond}|S_j^0] - p_{lj}^{\diamond})|S_j^0\big]$$

Then, performing the same series of operations as in the Proof of Proposition 1, we have

$$\begin{aligned}\mathbb{E}\big[(\mathbb{E}[p^{\diamond}_{lj}|S^0_j]-p^{\diamond}_{lj})'B(\mathbb{E}[p^{\diamond}_{lj}|S^0_j]-p^{\diamond}_{lj})|S^0_j\big] &= \mathbb{E}\big[(\mathbb{E}[\vec{x}|S^0_j]-\vec{x})'(\mathbf{e}^{(n+1)}_1+\mathbf{e}^{(n+1)}_{l+1})B(\mathbf{e}^{(n+1)}_1+\mathbf{e}^{(n+1)}_{l+1})'(\mathbb{E}[\vec{x}|S^0_j]-\vec{x})|S^0_j\big]\\ &= \mathbb{E}\big[(\mathbb{E}[\vec{x}|S^0_j]-\vec{x})'\tilde{\tilde{\Omega}}_{lj}(\mathbb{E}[\vec{x}|S^0_j]-\vec{x})|S^0_j\big]\\ &= tr\big(\tilde{\tilde{\Omega}}_{lj}\Sigma_{lj0}\big)\end{aligned}$$

where $\tilde{\tilde{\Omega}}_{lj} \equiv (\mathbf{e}^{(n+1)}_1+\mathbf{e}^{(n+1)}_{l+1})B(\mathbf{e}^{(n+1)}_1+\mathbf{e}^{(n+1)}_{l+1})'$ and it is indexed by l as it depends which regional division is solving the problem. Importantly, we can rewrite $\tilde{\tilde{\Omega}}_{lj}$ as

$$\begin{aligned}\tilde{\tilde{\Omega}}_l &\equiv (\mathbf{e}^{(n+1)}_1+\mathbf{e}^{(n+1)}_{l+1})B(\mathbf{e}^{(n+1)}_1+\mathbf{e}^{(n+1)}_{l+1})'\\ &= B\,\underbrace{(\mathbf{e}_1\mathbf{e}_1'+\mathbf{e}_1\mathbf{e}_{l+1}'+\mathbf{e}_{l+1}\mathbf{e}_1'+\mathbf{e}_{l+1}\mathbf{e}_{l+1}')}_{\Omega_l}\end{aligned}$$

where I drop the j index. We can then solve the following LGQ-RI

$$\begin{aligned}&\max_{\Sigma_{l,0}}-\frac{1}{2}\Big\{tr(\Sigma_{l,0}\Omega_l)-\omega\ln\Big(\frac{|\Sigma_{l,-1}|}{|\Sigma_{l,0}|}\Big)\Big\}\\ &\text{s.t. } \Sigma_{l,-1}-\Sigma_{l,0}\succeq 0\\ &\quad 0\prec\Sigma_{l,-1}\preceq\infty\end{aligned}$$

where objects are all indexed by l, as it is the regional division l who is solving the problem. Importantly, $\ln\Big(\frac{|\Sigma_{l,-1}|}{|\Sigma_{l,0}|}\Big)=\kappa_{lj}$, where κ_{lj} is now indexed by the regional division and the firm. I assume that $\Sigma_{l,-1}=\mathbf{I}$. The optimal posterior, $\Sigma_{l,0}$ is given by

$$\begin{aligned}\Sigma_{l,0} &= \omega_l\big[\text{Max}(\Omega_l,\omega_l)\big]^{-1}\\ &= \omega\mathbf{U}_l\max(\mathbf{D}_l,\omega)^{-1}\mathbf{U}_l^{-1}\end{aligned}$$

where $\Omega_l=\mathbf{U}_l\mathbf{D}_l\mathbf{U}_l$ is Ω_l eigendecomposition. First, note that we can write Ω_l as

$$\Omega_l=\begin{bmatrix}1 & \mathbf{e}_l'^{(1\times n)}\\ \mathbf{e}_l^{n\times 1} & \mathbf{e}_l^{(n\times 1)}\mathbf{e}_l'^{(1\times n)}\end{bmatrix}$$

To find its eigenvalues and eigenvectors, we want $det(\Omega_l - d\mathbf{I}) = 0$

$$det(\Omega_l - d\mathbf{I}) = \begin{bmatrix} 1-d & \mathbf{e}_l^{\prime(1\times n)} \\ \mathbf{e}_l^{n\times 1} & \mathbf{e}_l^{(n\times 1)}\mathbf{e}_l^{\prime(1\times n)} - d\mathbf{I}_n \end{bmatrix}$$

Since $(1-d)$ is invertible, we have

$$det(\Omega_l - d\mathbf{I}) = \begin{bmatrix} 1-d & \mathbf{e}_l^{\prime(1\times n)} \\ \mathbf{e}_l^{n\times 1} & \mathbf{e}_l^{(n\times 1)}\mathbf{e}_l^{\prime(1\times n)} - d\mathbf{I}_n \end{bmatrix} = det(1-d)\times det(\mathbf{e}_l^{(n\times 1)}\mathbf{e}_l^{\prime(1\times n)} - d\mathbf{I}_n - \mathbf{e}_l^{n\times 1}(1-d)^{-1}\mathbf{e}_l^{\prime(1\times n)})$$

$\mathbf{e}_l^{(n\times 1)}\mathbf{e}_l^{\prime(1\times n)} - d\mathbf{I}_n$ is a square diagonal matrix with non-zero main diagonal elements. Therefore, it is invertible. $-(1-d)^{-1}\mathbf{e}_l^{n\times 1}$ and $\mathbf{e}_l^{n\times 1}$ are column vectors. Therefore, we can use the matrix determinant lemma to calculate $det(\mathbf{e}_l^{(n\times 1)}\mathbf{e}_l^{\prime(1\times n)} - d\mathbf{I}_n - \mathbf{e}_l^{n\times 1}(1-d)^{-1}\mathbf{e}_l^{\prime(1\times n)})$. The resulting expression for $det(\Omega_l - d\mathbf{I})$ is

$$det(\Omega_l - d\mathbf{I}) = (2-\lambda)\times(-\lambda)^n = 0$$

Therefore, we have *only one* positive eigenvalue, with value equals to 2. The other n eigenvalues have zero value. When $d_1 = 2$, the eigenvector associated satisfies

$$-v_1 + v_{l+1} = 0, \quad 2v_i = 0, i \neq l+1$$

Hence, an eigenvector is $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1,0,\ldots,0,1,0,\ldots,0)' = \frac{1}{\sqrt{2}}(\mathbf{e}_1 + \mathbf{e}_{l+1}) \in \mathbb{R}^{n+1}$ and $\mathbf{D} = (d_{ii})_{i\{1,\ldots,n+1\}}$ such that $d_{ii} = 2, i = 1$ and $d_{ii} = 0, i > 1$. As a result, the regional division acquires exactly one signal about its *own* demand. The capacity constraint and the optimal posterior imply that

$$e^{-2\kappa_{lj}} = \omega_l^{n+1}|\max\left(\mathbf{D},\omega_l\right)^{-1}|$$

Since the capacity constraint is binding, we have that ω_l is non-negative. Furthermore, note that for the above equation to hold, ω_l cannot be zero. With positive ω_l, we have

$$\omega_l = 2e^{-2\kappa_{lj}}$$

Using Theorem 2.2 from Afrouzi and Yang (2021), we have that the regional division l from firm j acquires one signal about its *own* demand with the following structure:

$$s_{1lj}^{\mathrm{RD}} = \frac{1}{\sqrt{2}} p_l^{\diamond} + \nu_{1lj}, \quad \nu_{1lj} \sim \mathcal{N}\left(0, \frac{1}{e^{2\kappa_{lj}} - 1}\right)$$

and we have the result. □

B.2.2 Proof of Proposition 6

This result follow from Proposition 5 and Proposition 2.3 from Afrouzi and Yang (2021). Using them, we have

$$p_{lj}^{\mathrm{RD}} = \left(1 - e^{-2\kappa_{lj}}\right)\left[m + \lambda_l + \sqrt{2}\nu_{1lj}\right]$$

Hence,

$$\frac{\partial p_{lj}^{\mathrm{RD}}}{\partial m} = (1 - e^{-2\kappa_{lj}})$$

Assuming further that $\kappa_{lj} = \frac{\kappa}{k}$, we have

$$\frac{\partial\left(\partial p_{lj}^{\mathrm{RD}} / \partial m\right)}{\partial k} = -2\frac{\kappa}{k} e^{-2\frac{\kappa}{k}} < 0$$

□

B.2.3 Proof of Corollary 1

To see that, we need to compare $\frac{\partial p_{lj}^{\mathrm{HQ}}}{\partial m}, \frac{\partial p_{lj}^{\mathrm{RD}}}{\partial m}$. When, $k > e^{2\kappa} - 1$, it is clear, as $\kappa > \frac{\kappa}{k}$. For $k \leq e^{2\kappa} - 1$,

$$\begin{aligned}
\frac{\partial p_{lj}^{\mathrm{HQ}}}{\partial m} > \frac{\partial p_{lj}^{RD}}{\partial m} &\iff 1 - \frac{1}{(k+1)^{1-\frac{1}{k}}(e^{2\kappa})^{\frac{1}{k}}} > 1 - e^{-2\frac{\kappa}{k}} \\
&\iff \frac{1}{(k+1)^{1-\frac{1}{k}}(e^{2\kappa})^{\frac{1}{k}}} < e^{-2\frac{\kappa}{k}} \\
&\iff \frac{1}{(k+1)^{1-\frac{1}{k}}} < 1 \\
&\iff 1 < (k+1)^{1-\frac{1}{k}}
\end{aligned}$$

which holds for any $k \geq 1$. $\square$

B.2.4 Proof of Proposition 7

This result follow from Proposition 5 and Proposition 2.3 from Afrouzi and Yang (2021). Using them, we have

$$p_{lj}^{\mathrm{RD}} = \left(1 - e^{-2\kappa_{lj}}\right)[m + \lambda_l + \sqrt{2}\nu_{1lj}]$$

Hence

$$\frac{\partial p_{lj}^{\mathrm{RD}}}{\partial \lambda_l} = \left(1 - e^{-2\kappa_{lj}}\right) \quad \text{and} \quad \frac{\partial p_{lj}^{\mathrm{RD}}}{\partial \lambda_\ell} = 0, \quad \ell \neq l$$

$\square$

B.2.5 Proof of Corollary 2

This is a direct result of Proposition 4 and Proposition 7.

B.3 Additional Results and Extensions of the Static Model

B.3.1 Proof of Proposition A.1

This result follows from the signal structure in Proposition 1 and Proposition 2.3 from Afrouzi and Yang (2021). $\square$

B.3.2 Proof of Proposition A.2

The outline of the proof of this result is very similar to Proposition 1, except that instead of κ, we consider $\kappa(k) : \mathbb{R}_+ \mapsto \mathbb{R}_+$. In particular, the capacity constraint becomes $|\Sigma_0| = e^{-2\kappa(k)}$. After calculating the eigenvalues of $\Omega(k)$, we have two cases: (1) $\omega \in (1/k, 1 + 1/k]$, which implies in one eigenvalue larger than ω; (2) $\omega \in (0, 1/k]$ which implies in k eigenvalues larger than ω.

Case 1: $\omega \in (1/k, 1 + 1/k]$. In this case, the firm acquires one only signal. Using the capacity constaint and the optimal posterior, we have

$$\omega(k) = \left(1 + \frac{1}{k}\right)e^{-2\kappa(k)}$$

So we must have

$$\left(1+\frac{1}{k}\right)e^{-2\kappa(k)} > \frac{1}{k} \iff \frac{1}{2}\ln(1+k) > \kappa(k)$$

Also, we must have

$$\left(1+\frac{1}{k}\right)e^{-2\kappa(k)} \leq \left(1+\frac{1}{k}\right) \iff e^{2\kappa(k)} \geq 1$$

Case 2: $\omega \in (0, 1/k]$**.** In this case, the firm acquires k signals. Using the capacity constaint and the optimal posterior, we have

$$\omega(k) = \frac{1}{k}\left[(k+1)\frac{1}{e^{2\kappa(k)}}\right]^{\frac{1}{k}}$$

So we must have

$$\frac{1}{k}\left[(k+1)\frac{1}{e^{2\kappa(k)}}\right]^{\frac{1}{k}} \leq \frac{1}{k} \iff \frac{1}{2}\ln(k+1) \leq \kappa(k)$$

By assumption, we have that $\exists \bar{k} > 0$, such that $\frac{1}{2}\ln(\bar{k}+1) = \kappa(\bar{k})$, $\frac{1}{2}\ln(1+k) < \kappa(k)$ for $k < \bar{k}$, and $\frac{1}{2}\ln(1+k) > \bar{k}$. This implies that for $k < \bar{k}$, the firm acquires k signals and for $k > \bar{k}$, the firm acquires only one signal, which is the first part of the proposition. For the signal structure, note that the $\Omega(k)$ is the same as in Proposition 1. Therefore, the signal loadings are going to be given by the eigenvectors associated with the eigenvalues of $\Omega(k)$ as in Proposition 1. □

B.3.3 Proof of Proposition A.3

The outline of the proof of this result is very similar to Proposition 1, except that the prior uncertainty is not equal to the identity matrix. This changes the optimal posterior. Yet, under the assumption I made about the Σ_{-1}, the optimal posterior is still very tractable. When we assume

$$\Sigma_{-1} = \begin{bmatrix} \sigma_m^2 & \mathbf{0}'_{1\times n} \\ \mathbf{0}_{n\times 1} & \sigma_\lambda^2 \mathbf{I}_{n\times n} \end{bmatrix}$$

the capacity constraint becomes

$$\begin{aligned}|\Sigma_0| &= |\Sigma_{-1}|e^{-2\kappa}\\ &= \sigma_m^2 \times (\sigma_\lambda^2)^n e^{-2\kappa}\end{aligned}$$

and the optimal posterior

$$\Sigma_0 = \omega\Sigma_{-1}^{1/2}\left[\mathrm{Max}\left(\Sigma_{-1}^{1/2}\Omega(k)\Sigma_{-1}^{1/2},\omega\right)\right]^{-1}\Sigma_{-1}^{1/2}$$

In particular, we need to calculate the eigenvalues of

$$\Sigma_{-1}^{1/2}\Omega(k)\Sigma_{-1}^{1/2} = \begin{bmatrix} \sigma_m^2 & \frac{1}{k}\sigma_m\sigma_\lambda\mathbf{1}'_{1\times k} & \mathbf{0}_{1\times(n-k)} \\ \frac{1}{k}\sigma_m\sigma_\lambda\mathbf{1}_{k\times 1} & \frac{1}{k}\sigma_\lambda^2\mathbf{I}_{k\times k} & \mathbf{0}_{k\times(n-k)} \\ \mathbf{0}_{(n-k)\times 1} & \mathbf{0}_{(n-k)\times k} & \mathbf{0}_{(n-k)\times(n-k)} \end{bmatrix}$$

Using the expression of determinant of block matrices and the matrix determinant lemma as I did in Proposition 1, I get the following:

$$det(\Sigma_{-1}^{1/2}\Omega(k)\Sigma_{-1}^{1/2} - d\mathbf{I}) = \left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2 - d\right) \times (\frac{1}{k}\sigma_\lambda^2 - d)^{k-1} \times (-d)^{n-k+1} = 0$$

Therefore, there's one eigenvalue with value $\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)$ and $k-1$ eigenvalues with value $\frac{1}{k}\sigma_\lambda^2$. In particular, if we set $\sigma_m^2 = \sigma_\lambda^2 = 1$, we are back to the Proposition 1. Let's find the eigenvectors associated with each one of these eigenvalues. First, for $d_1 = \left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)$, we have the following

$$\begin{bmatrix} \sigma_m^2 - \sigma_m^2 - \frac{1}{k}\sigma_\lambda^2 & \frac{1}{k}\sigma_m\sigma_\lambda\mathbf{1}'_{1\times k} & \mathbf{0}_{1\times(n-k)} \\ \frac{1}{k}\sigma_m\sigma_\lambda\mathbf{1}_{k\times 1} & \frac{1}{k}\sigma_\lambda^2\mathbf{I}_{k\times k} - \left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)\mathbf{I} & \mathbf{0}_{k\times(n-k)} \\ \mathbf{0}_{(n-k)\times 1} & \mathbf{0}_{(n-k)\times k} & -\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)\mathbf{I}_{(n-k)\times(n-k)} \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

implying in

$$-\sigma_\lambda v_1 + \sum_{j=2}^{k+1} \sigma_m v_j = 0, \quad \frac{1}{k}\sigma_\lambda v_1 - \sigma_m v_j = 0, j = 2, \dots, k+1, \quad -\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right) v_j = 0, j = k+2, \dots, n+1$$

Therefore, an eigenvector associated with $d_1 = \left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)$ is

$$\mathbf{u}_1 = \frac{1}{\sqrt{1 + \frac{1}{k}\frac{\sigma_\lambda^2}{\sigma_m^2}}} \times \left(1, \frac{\sigma_\lambda}{\sigma_m}\frac{1}{k}, \dots, \frac{\sigma_\lambda}{\sigma_m}\frac{1}{k}, 0, \dots 0\right)'$$

For $d_2 = \frac{1}{k}\sigma_\lambda^2$, we have

$$\begin{bmatrix} \sigma_m^2 - \frac{1}{k}\sigma_\lambda^2 & \frac{1}{k}\sigma_m\sigma_\lambda \mathbf{1}'_{1\times k} & \mathbf{0}_{1\times(n-k)} \\ \frac{1}{k}\sigma_m\sigma_\lambda \mathbf{1}_{k\times 1} & \frac{1}{k}\sigma_\lambda^2 \mathbf{I}_{k\times k} - \left(\frac{1}{k}\sigma_\lambda^2\right)\mathbf{I} & \mathbf{0}_{k\times(n-k)} \\ \mathbf{0}_{(n-k)\times 1} & \mathbf{0}_{(n-k)\times k} & -\left(\frac{1}{k}\sigma_\lambda^2\right)\mathbf{I}_{(n-k)\times(n-k)} \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and

$$v_1 = 0, \quad \sum_{j=2}^{k+1} v_j = 0, \quad v_j = 0, j = k+2, \dots n+1$$

Now, we can find the optimal posterior. There are two cases: (1) $\omega \in (\frac{1}{k}\sigma_\lambda^2, \sigma_m^2 + \frac{1}{k}\sigma_\lambda^2]$; (2) $\omega \in (0, \frac{1}{k}\sigma_\lambda^2]$.

Case 1: $\omega \in (\frac{1}{k}\sigma_\lambda^2, \sigma_m^2 + \frac{1}{k}\sigma_\lambda^2]$. In this case, the firm acquires one signal. Using the capacity constraint and the optimality conditional for optimal posterior, we get:

$$\omega(\kappa, k, \sigma_m, \sigma_\lambda) = \left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right) e^{-2\kappa}$$

and we must have

$$\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right) e^{-2\kappa} > \frac{1}{k}\sigma_\lambda^2 \iff k > (e^{2\kappa} - 1)\frac{\sigma_\lambda^2}{\sigma_m^2}$$

Case 2: $\omega \in (0, \frac{1}{k}\sigma_\lambda^2]$**.** In this case, the firm acquires k signals. Using the capacity constraint and the optimality conditional for optimal posterior, we get:

$$\omega(\kappa, k, \sigma_m, \sigma_\lambda) = \left(\frac{1}{k}\sigma_\lambda^2\right)^{\frac{k-1}{k}} \times \left[\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)e^{-2\kappa}\right]^{\frac{1}{k}}$$

and we must have

$$\left(\frac{1}{k}\sigma_\lambda^2\right)^{1-\frac{1}{k}} \times \left[\left(\sigma_m^2 + \frac{1}{k}\sigma_\lambda^2\right)e^{-2\kappa}\right]^{\frac{1}{k}} \leq \frac{1}{k}\sigma_\lambda^2 \iff k \leq (e^{2\kappa} - 1)\frac{\sigma_\lambda^2}{\sigma_m^2}$$

Therefore, we have a similar result to Proposition 1. There is threshold of number of regions $\underline{k}(\kappa, \sigma_m, \sigma_\lambda) = (e^{2\kappa} - 1)\frac{\sigma_\lambda^2}{\sigma_m^2}$ such that if firm operates in more regions, it acquires a single signal. If it operates in fewer regions, it acquires k signals. The distinguishing feature is that now the relative prior uncertainty about the shocks also affect this threshold, in addition to the capacity. Finally, using Theorem 2.2 from Afrouzi and Yang (2021), I can find the signal structure in each one of these cases. □

C Equilibrium Definition

In this section, I define the rational inattention, the flexible price, and the non-stochastic efficient steady state equilibria.

Definition 1. A **rational inattention** equilibrium is a set of allocations for the households $\mathscr{A}_h(l) = \{(C_{ljt})_{j\in[0,1]}, L_{lt}^s, B_{lt}\}_{t\geq 0}$ in each region $l \in [n]$, an allocation for all firms $\mathscr{A}_f = \{(Y_{ljt}^s)_{l\in\mathscr{L}_j}, (L_{ljt})_{l\in\mathscr{L}_j}, S_{jt}, (P_{ljt})_{l\in\mathscr{L}_j}\}_{j\in[0,1],t\geq 0}$; a set of monetary and fiscal policies $\mathscr{A}_g = \{M_t, (\tau_{lt})_{l\in[n]}, (T_{lt})_{l\in[n]}\}_{t\geq 0}$; and a set of prices $\mathscr{P} = \{(P_{lt})_{l\in[n]}, W_t, i_t\}_{t\geq 0}$ such that given the processes for regional markups $\{\Lambda_{lt}\}_{l\in[n],t\geq 0}$:

1. Given $\{\Lambda_{lt}\}_{l\in[n],t\geq 0}$, $\mathscr{P}$, and $\mathscr{A}_g$, $\mathscr{A}_h(l)$ solves the household's problem in Equations (18) to (20), $\forall\, l \in [n]$

2. Given $\{\Lambda_{lt}\}_{l\in[n],t\geq 0}$, $\mathscr{P}$, and $\mathscr{A}_g$, $\mathscr{A}_f$ solves the monopolistic competitive firm's problem in Equations (22) to (25), $\forall\, j \in [0,1]$

3. Labor, bond holdings, and retailers' regional good markets clear and government budget constraint is satified:

   (a) $\sum_{l\in[n]} \zeta_l L_{lt}^s = \int_0^1 \sum_{l\in\mathscr{L}_j} L_{ljt} dj$

(b) $\sum_{l\in[n]} \zeta_l B_{lt} = 0$

(c) $C_{ljt} = Y^s_{ljt}, \forall l \in [n], j \in [0,1]$

(d) $\sum_l T_{lt} = \int_0^1 \sum_{l\in\mathscr{L}_j} \tau_{lt} P_{ljt} Y_{ljt} dj$

Definition 2. A **flexible price** equilibrium is a set of allocations for the households $\mathscr{A}_h(l) = \{(C_{ljt})_{j\in[0,1]}, L^s_{lt}, B_{lt}\}_{t\geq 0}$ in each region $l \in [n]$, an allocation for all firms $\mathscr{A}_f = \{(Y^s_{ljt})_{l\in\mathscr{L}_j}, (L_{ljt})_{l\in\mathscr{L}_j}, (P_{ljt})_{l\in\mathscr{L}_j}\}_{j\in[0,1],t\geq 0}$; a set of monetary and fiscal policies $\mathscr{A}_g = \{M_t, (\tau_{lt})_{l\in[n]}, (T_{lt})_{l\in[n]}\}_{t\geq 0}$; and a set of prices $\mathscr{P} = \{(P_{lt})_{l\in[n]}, W_t, i_t\}_{t\geq 0}$ such that given the processes for regional markups $\{\Lambda_{lt}\}_{l\in[n],t\geq 0}$, and initial set of signals $\{S^{-1}_j\}_{j\in[0,1]}$:

1. Given $\{\Lambda_{lt}\}_{l\in[n],t\geq 0}$, $\mathscr{P}$, and $\mathscr{A}_g$, $\mathscr{A}_h(l)$ solves the household's problem in Equations (18) to (20), $\forall l \in [n]$

2. Given $\{\Lambda_{lt}\}_{l\in[n],t\geq 0}$, $\mathscr{P}$, and $\mathscr{A}_g$, $\mathscr{A}_f$ solves the monopolistic competitive firm's problem in Equation (21), $\forall j \in [0,1]$

3. Labor, bond holdings, and retailers' regional good markets clear and government budget constraint is satified:

    (a) $\sum_{l\in[n]} \zeta_l L^s_{lt} = \int_0^1 \sum_{l\in\mathscr{L}_j} L_{ljt} dj$

    (b) $\sum_{l\in[n]} \zeta_l B_{lt} = 0$

    (c) $C_{ljt} = Y^s_{ljt}, \forall l \in [n], j \in [0,1]$

    (d) $\sum_l T_{lt} = \int_0^1 \sum_{l\in\mathscr{L}_j} \tau_{lt} P_{ljt} Y_{ljt} dj$

Definition 3. A **non-stochastic efficient steady state** equilibrium is a set of *time-invariant* allocations for the households $\mathscr{A}_h(l) = \{(C_{lj})_{j\in[0,1]}, L^s_l, B_l\}$ in each region $l \in [n]$, a *time-invariant* allocation for all firms $\mathscr{A}_f = \{(Y^s_{lj})_{l\in\mathscr{L}_j}, (L_{lj})_{l\in\mathscr{L}_j}\}_{j\in[0,1],(P_{lj})_{l\in\mathscr{L}_j}}$; a set of *time-invariant* monetary and fiscal policies $\{M, (\tau_l)_{l\in[n]}, (T_l)_{l\in[n]}\}$; and a set of *time-invariant* prices $\mathscr{P} = \{(P_l)_{l\in[n]}, W, i\}$ such that given *time-invariant* regional markups $\{\Lambda_l\}_{l\in[n]}$:

1. Given $\{\Lambda_l\}_{l\in[n]}$, $\mathscr{P}$, and $\mathscr{A}_g$, $\mathscr{A}_h(l)$ solves the household's problem in Equations (18) to (20), $\forall l \in [n]$

2. Given $\{\Lambda_l\}_{l\in[n]}$, $\mathscr{P}$, and $\mathscr{A}_g$, $\mathscr{A}_f$ solves the monopolistic competitive firm's problem in Equation (21), $\forall j \in [0,1]$

3. Labor, bond holdings, and retailers' regional good markets clear and government budget constraint is satified:

   (a) $\sum_{l\in[n]} \zeta_l L^s_l = \int_0^1 \sum_{l\in\mathscr{L}_j} L_{lj} dj$

   (b) $\sum_{l\in[n]} \zeta_l B_l = 0$

   (c) $C_{lj} = Y^s_{lj}, \forall l \in [n], j \in [0,1]$

   (d) $\sum_l T_l = \int_0^1 \sum_{l\in\mathscr{L}_j} \tau_l P_{lj} Y_{lj} dj$

   with taxes that undo distortions from monopolistic competition, $\tau_l = 1-\Lambda_l, \forall l \in [n]$.

D Derivations of Optimality Conditions in the Model

Here, I characterize the flexible-price, imperfect information, and rational inattention optimality conditions of the economy. The only block that changes is the supply side of the economy.

D.1 Household's Optimality Conditions

Consider a representative household in region $l \in [n]$. First, we solve for the expenditure minimization problem, given a level of consumption C_{lt} and given a vector of retailer-location prices $\{P_{ljt}\}_{j\in[0,1]}$. That is,

$$\mathscr{E}(C_{lt}; \{P_{ljt}\}_{j\in[0,1]}) = \min_{\{C_{ljt}\}_{j\in[0,1]}} \int_0^1 P_{ljt} C_{ljt} dj \quad \text{subject to } C_{lt} = \Big(\int_0^1 \theta_{lj}^{\frac{\Lambda_{lt}-1}{\Lambda_{lt}}} C_{ljt}^{\frac{1}{\Lambda_{lt}}} dj\Big)^{\Lambda_{lt}} \geq C_{lt} \tag{65}$$

After defining the Lagrange multiplier associated with the constraint of the problem as the regional price index, P_{lt}, the household l demand for retailer good j in period t is given by

$$C_{ljt} = \theta_{lj} \left(\frac{P_{ljt}}{P_{lt}}\right)^{\frac{\Lambda_{lt}}{1-\Lambda_{lt}}} C_{lt} \tag{66}$$

Note that this is a demand *per capita*. The regional price index is given by

$$P_{lt} = \left(\int_0^1 \theta_{lj} P_{ljt}^{\frac{1}{1-\Lambda_{lt}}} dj\right)^{1-\Lambda_{lt}} \tag{67}$$

with $P_{lt}C_{lt} = \int_0^1 P_{ljt}C_{ljt}dj$. The share of expenditure in location l on retailer j is given by:

$$\frac{P_{ljt}C_{ljt}}{P_{lt}C_{lt}} = \theta_{lj}\left(\frac{P_{ljt}}{P_{lt}}\right)^{\frac{1}{1-\Lambda_{lt}}} \tag{68}$$

Given the processes for prices $\{P_{lt}, W_t\}_{t\geq 0}$, the household maximizes

$$\max_{\{C_{lt},L_{lt}\}_{t\geq 0}} \mathbb{E}_0^f\Big[\sum_{t=0}^{\infty}\beta^t\big(\log(C_{lt}) - L_{lt}\big)\Big]$$
$$\text{s.t.} \quad P_{lt}C_{lt} + B_{lt} \leq W_t L_{lt} + (1+i_{t-1})B_{lt-1} + \text{Profits}_{lt} + T_{lt}$$

Let μ_t^l be the lagrange multiplier associated with the household's l budget constraint at time t. The first order conditions imply in:

$$\beta^t\frac{1}{P_{lt}C_{lt}} = \mu_t^l \tag{69}$$

$$\beta^t\frac{1}{W_t} = \mu_t^l \tag{70}$$

$$\mu_{lt} = (1+i_t)\mathbb{E}\big[\mu_{lt+1}\big] \tag{71}$$

Combining Equation (69) and Equation (70), we get that the nominal GDP in region l is equal to economy-wide wages:

$$P_{lt}C_{lt} = W_t \tag{72}$$

Combining Equation (70) and Equation (71), we get that the region l euler equation is given by

$$\beta(1+i_t)\mathbb{E}\Big[\frac{W_t}{W_{t+1}}\Big] = 1 \tag{73}$$

Since W_t is the same across regions due to the single labor market assumption, Equation (73) implies that the *regional* Euler equation *does not* depend on regional variables.

D.2 Firms' Optimality Conditions

Consider a retailer $j \in [0,1]$, let $\mathscr{L}_j$ be the set of region in which j is present. First, we solve for the firm's cost minimization problem in region $l \in \mathscr{L}_j$, given Y_{ljt} and W_t. That is,

$$\mathscr{C}_{ljt}(Y_{ljt}; W_t) = \min_{L_{ljt}} W_t L_{ljt} \quad \text{subject to } L_{ljt} \geq Y_{ljt} \tag{74}$$

Equation (74) implies that firm's marginal cost is given by

$$\mathsf{MC}_{ljt}(W_t) = W_t \tag{75}$$

Flexible Pricing

Under full information, firm j maximization's problem becomes static. That is,

$$\max_{(P_{ljt})_{l\in\mathscr{L}_j}} \sum_{l\in\mathscr{L}_j} ((1-\tau_{lt})P_{ljt} - W_t) \times Y^s_{ljt} \quad \text{subject to } Y^s_{ljt} = \zeta_l C_{ljt} = \zeta_l \mathscr{D}(P_{ljt}/P_{lt}; C_{lt}) \tag{76}$$

Furthermore, since firm's demand in region l depends only on its relative prices in region l, we can solve for the firm's optimal pricing in each region $l \in \mathscr{L}_j$ separately. Given the CES demand function in Equation (66), the optimal pricing in the absence of any friction is given by

$$P^{\diamond}_{ljt} = \frac{1}{(1-\tau_{lt})}\Lambda_{lt} \times \mathsf{MC}_{ljt} = \frac{1}{(1-\tau_{lt})}\Lambda_{lt} \times W_t \tag{77}$$

Equation (77) characterizes the desired price firm j wants to set in region l at time t.

Optimal Pricing under Imperfect Information

Given firm's marginal cost in region l, the firm's optimal pricing problem *given* a sequence of signals $\{S_t\}_{t\geq 0}$ is given by

$$(P^*_{ljt})_{l\in\mathscr{L}_j} = \text{argmax}_{(P_{ljt})_{l\in\mathscr{L}_j}} \mathbb{E}\Big[\sum_{l\in\mathscr{L}_j} ((1-\tau_{lt})P_{ljt} - W_t) \times \zeta_l \mathscr{D}(P_{ljt}/P_{lt}; C_{lt}) \Big| S^t \Big] \tag{78}$$

Optimal Pricing under Rational Inattention

Under Rational Inattention, given a capacity κ, the firms solves the following problem:

$$\max_{\{S_{j,t}\subset\mathbb{S}^t,\{P_{ljt}(S_j^t)\}_{l\in\mathscr{L}_j}\}_{t\geq 0}} \mathbb{E}\Big[\sum_{t=0}^{\infty}\underbrace{\beta^t W_t^{-1}}_{\substack{\text{discount}\\ \text{factor}}}\times\Big\{\sum_{l\in\mathscr{L}_j}\Big(\underbrace{(1-\tau_l)P_{ljt}Y_{ljt}^s}_{\substack{\text{revenue}\\ \text{in } l}}-\underbrace{W_t L_{ljt}}_{\substack{\text{production}\\ \text{cost in } l}}\Big)\Big\}\Big] \tag{79}$$

$$\text{s.t.}\quad Y_{ljt}^s=\zeta_l C_{ljt},\, l\in\mathscr{L}_j \qquad \text{(demand)} \tag{80}$$

$$\mathbb{I}(S_j^t;\vec{x}^t|S_j^{t-1})\leq\kappa_j \qquad \text{(info. processing constraint)} \tag{81}$$

$$S_j^t=S_j^{t-1}\cup S_{jt},\quad S_j^{-1}\text{ given} \qquad \text{(evolution of information set)} \tag{82}$$

where τ_l is a constant tax to firms in location l that eliminates steady-state inefficiencies coming from monopolistic competition. From now on, I assume that $\kappa_j=\kappa, \forall j\in[0,1]$. That is, all firms have the same capacity. S_j^{-1} is an initial signal. Finally, S_j^t satisfies the no-forgetting condition, which states that firms do not forget information over time. This will put an upper bound on the amount of uncertainty the firm can choose. This part is the same as in the main text and put here for completeness.

E Non-Stochastic Efficient Steady State

I log-linearize the model around a non-stochastic efficient steady state where the Central Bank targets a constant nominal aggregate GDP, normalized to one, $M=1$, and fiscal policy that undoes distortions from monopolistic competition in each region $l\in[n]$, $\tau_l=1-\Lambda_l$. In the non-stochastic efficient steady state, region l household's first order conditions in Equation (72) and Equation (73) imply in:

$$P_l C_l = W \tag{83}$$

The constant nominal aggregate GDP targeting implies in

$$M\equiv PC=\prod_{l\in[n]}(P_l C_l)^{\zeta_l} \tag{84}$$

Since $P_l C_l = W$, we have

$$M = W \tag{85}$$

In the non-stochastic efficient steady state, the optimal flexible pricing implies in

$$P^{\diamond}_{lj} = W \tag{86}$$

Plugging this into the regional price index, we get

$$\begin{aligned} P_l &= \Big(\int_0^1 \theta_{lj} W^{\frac{1}{1-\Lambda_l}} dj\Big)^{1-\Lambda_l} \\ &= W\Big(\int_0^1 \theta_{lj} dj\Big)^{1-\Lambda_l} \end{aligned} \tag{87}$$

Now, using Equation (85) with $M = 1$,

$$P_l = \Big(\int_0^1 \theta_{lj} dj\Big)^{1-\Lambda_l} \tag{88}$$

The household l's consumption expenditure share on retailer j will be given by

$$\frac{P_{lj} C_{lj}}{P_l C_l} = \frac{\theta_{lj}}{\int_0^1 \theta_{lj} dj} \tag{89}$$

Normalizing $\int_0^1 \theta_{lj} dj = 1, \forall l \in [n]$[14], we get

$$P_l = 1 \tag{90}$$

and

$$\frac{P_{lj} C_{lj}}{P_l C_l} = \theta_{lj} \tag{91}$$

[14]This is not an innocuous assumption, as it makes the steady-state regional price level to be the same across regions. As a result, the regional GDP is the same across regions in the non-stochastic efficient steady state.

Equation (91) will be used to calibrate $\{\theta_{lj}\}_{l\in[n],j\in[0,1]}$. Finally, plugging Equation (90) into Equation (83), using $M=1$ we get

$$C_l = \frac{W}{P_l} = 1 \tag{92}$$

An implication of the non-stochastic efficient steady state is that $P^{\diamond}_{lj} = 1, \forall l \in [n], j \in [0,1]$. That is, in the non-stochastic efficient steady state, firm j sets a uniform price across regions where it is located, even with heterogeneous taste shifters and heterogeneous regional markups. The single labor market assumption, along with the absence of region-specific productivity shocks are key for this result. It will be useful for the calibration and second order approximation to know the non-stochastic efficient steady state firm j sales in a given region $l \in \mathcal{L}_j$. This is given by

$$\begin{aligned}
\mathsf{sales}_{lj} = (1-\tau_l)P^{\diamond}_{lj}Y^s_{ljt} &= P^{\diamond}_{lj}\zeta_l\theta_{lj}\left(\frac{P^{\diamond}_{lj}}{P_l}\right)^{\frac{\Lambda_l}{1-\Lambda_l}}C_l \\
&= (1-\tau_l)\zeta_l\theta_{lj}\frac{P^{\diamond}_{lj}}{P_l}\left(\frac{P^{\diamond}_{lj}}{P_l}\right)^{\frac{\Lambda_l}{1-\Lambda_l}}W \\
&= (1-\tau_l)\zeta_l\theta_{lj}\left(\frac{P^{\diamond}_{lj}}{P_l}\right)^{\frac{1}{1-\Lambda_l}} \\
&= \Lambda_l\zeta_l\theta_{lj}
\end{aligned}$$

Firm j's total sales are given by

$$\mathsf{sales}_j = \sum_{l\in[n]} \Lambda_l\zeta_l\theta_{lj} \tag{93}$$

F Log-Linearization

This section derives the log-linearized optimality conditions. Let small letters denote the log deviations of their corresponding variables from their non-stochastic steady state equilibrium values. That is, $x_t \equiv \ln(X_t/X)$.

Region l's price index. From Equation (67), we have

$$\begin{aligned}
P_{lt}^{\frac{1}{1-\Lambda_{lt}}} - \int_0^1 \theta_{lj}P_{ljt}^{\frac{1}{1-\Lambda_{lt}}}\,dj = 0 \\
\left(P_l e^{p_{lt}}\right)^{\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}} - \int_0^1 \theta_{lj}\left(P_{lj}e^{p_{ljt}}\right)^{\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}}\,dj = 0
\end{aligned}$$

Let $f(p_{lt}, \lambda_{lt}, (p_{ljt})_{j\in[0,1]}) \equiv \left(P_l e^{p_{lt}}\right)^{\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}} - \int_0^1 \theta_{lj}\left(P_{lj} e^{p_{ljt}}\right)^{\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}} dj$. Performing a first-order Taylor expansion of $f(.)$ around $\mathbf{0}$ gives us

$$f(p_{lt}, \lambda_{lt}, (p_{ljt})_{j\in[0,1]}) \approx f(\mathbf{0}) + f_{p_{lt}}(\mathbf{0})(p_{lt} - 0) + f_{\lambda_{lt}}(\mathbf{0})(\lambda_{lt} - 0) + \int_0^1 f_{p_{ljt}}(\mathbf{0})(p_{ljt} - 0) dj$$

where $f(\mathbf{0}) = 0$, $f_{p_{lt}}(\mathbf{0}) = \frac{1}{1-\Lambda_l} P_l^{\frac{1}{1-\Lambda_l}}$, $f_{\lambda_{lt}}(\mathbf{0}) = 0$, $f_{p_{ljt}}(\mathbf{0}) = -\theta_{lj} \frac{1}{1-\Lambda_l} P_{lj}^{\frac{1}{1-\Lambda_l}}$, which gives us

$$\begin{aligned} f(p_{lt}, \lambda_{lt}, (p_{ljt})_{j\in[0,1]}) &\approx \frac{1}{1-\Lambda_l} P_l^{\frac{1}{1-\Lambda_l}} p_{lt} + \int_0^1 (-\theta_{lj} \frac{1}{1-\Lambda_l} P_{lj}^{\frac{1}{1-\Lambda_l}}) p_{ljt} dj \\ &\approx \frac{1}{1-\Lambda_l}(1-\beta)^{\frac{1}{1-\Lambda_l}} p_{lt} - \frac{1}{1-\Lambda_l}(1-\beta)^{\frac{1}{1-\Lambda_l}} \int_0^1 \theta_{lj} p_{ljt} dj \end{aligned}$$

Hence, up to a first order we have

$$p_{lt} = \int_0^1 \theta_{lj} p_{ljt} dj \tag{94}$$

Household in region l optimal intratemporal condition. Taking logs in Equation (72), we get

$$p_{lt} + c_{lt} = w_t \tag{95}$$

Household in region l optimal intertemporal condition. Let $R_t \equiv (1 + i_t)$. We can rewrite Equation (73) as

$$\begin{aligned} \mathbb{E}\left[\beta R_t \frac{W_t}{W_{t+1}}\right] &= 1 \\ \mathbb{E}\left[\beta R e^{r_t} \frac{e^{w_t}}{e^{w_{t+1}}}\right] &= 1 \end{aligned}$$

Now, let $f(r_t, w_t, w_{t+1}) = \beta R e^{r_t} \frac{e^{w_t}}{e^{w_{t+1}}}$. Performing a first-order Taylor expansion of $f(.)$ around $(0,0,0)$ gives us

$$f(r_t, w_t, w_{t+1}) \approx f(0,0,0) + f_{r_t}(0,0,0)(r_t - 0) + f_{w_t}(0,0,0)(w_t - 0) + f_{w_{t+1}}(0,0,0)(w_{t+1} - 0)$$

where $f(0,0,0) = \beta R$, $f_{r_t}(0,0,0) = \beta R$, $f_{w_t}(0,0,0) = \beta R$, and $f_{w_{t+1}}(0,0,0) = -\beta R$. Hence,

$$f(r_t, w_t, w_{t+1}) \approx \beta R + \beta R r_t + \beta R w_t - \beta R w_{t+1}$$

Recall that $\beta R = 1$. This implies that up to a first-order we have

$$r_t + w_t - \mathbb{E}\left[w_{t+1}\right] = 0 \tag{96}$$

Monopolistic competitive retailer's marginal cost. Taking logs in Equation (75), we get

$$mc_{ljt} = w_t \tag{97}$$

Ideal prices. Taking logs of Equation (77) and recalling there's no disturbances to regional taxes, we get

$$p^{\diamond}_{ljt} = \lambda_{lt} \times mc_{ljt} = \lambda_{lt} \times w_t \tag{98}$$

Aggregate real GDP and aggregate prices. From the definitions of aggregate real GDP and aggregate prices, we have

$$c_t = \sum_{l \in [n]} \zeta_l c_{lt} \tag{99}$$

$$p_t = \sum_{l \in [n]} \zeta_l p_{lt} \tag{100}$$

$$m_t \equiv p_t + c_t = \sum_{l \in [n]} \zeta_l (p_{lt} + c_{lt}) \tag{101}$$

Wages. To characterize the log-linearization of wages, we combine Equation (95) and Equation (101)

$$\begin{aligned} m_t &= \sum_{l \in [n]} \zeta_l (p_{lt} + c_{lt}) \\ &= \sum_{l \in [n]} \zeta_l (w_t) \\ m_t &= w_t \end{aligned} \tag{102}$$

Interest rates. Finally, to characterize the interest rates r_t, we combine Equation (102), Equation (96), and the process for nominal GDP $m_t = m_{t-1} + \sigma_u u_t, u_t \sim N(0,1)$

$$
\begin{aligned}
r_t + w_t - \mathbb{E}[w_{t+1}] &= 0 \\
r_t + m_t - \mathbb{E}[m_{t+1}] &= 0 \\
r_t + m_t - \mathbb{E}[m_t + \sigma_u u_{t+1}] &= 0
\end{aligned}
$$

which implies in

$$r_t = 0 \tag{103}$$

and, using the definition of $r_t \equiv \log \frac{(1+i_t)}{(1+i)}$

$$i_t = \frac{1}{\beta} - 1 \tag{104}$$

where we use that in the steady state $(1+i) = 1/\beta$. Therefore, note that under nominal GDP targeting, the interest rates are *endogenously* constant and equal to $1/\beta - 1$.

Fundamental shocks. The shocks $\{m_t, \lambda_{1t}, \dots, \lambda_{nt}\}$ can be characterized as

$$\vec{x}_t = \mathbf{A}\vec{x}_{t-1} + \mathbf{Q}\vec{u}_t, \vec{u}_t \sim \mathcal{N}(\vec{0}, \mathbf{I}) \tag{105}$$

G Second Order Approximation of the Profit Function

I solve the firm's rational inattention problem by performing a second order approximation of the profit function around the non-stochastic efficient steady state. The derivation of the profit function follows Mackowiak and Wiederholt (2009, 2015). Let the profit of the firm j be given by

$$
\begin{aligned}
\Pi_j((P_{lj})_{l\in[\mathscr{L}_j]}; W_t, (P_{lt})_{l\in[n]}, (\Lambda_{lt})_{l\in[n]}) &= \sum_{l=1}^{n} \zeta_l \frac{U'(C_{lt})}{P_{lt}} \left[\sum_{l\in[\mathscr{L}_j]} \zeta_l \theta_{lj} \left((1-\tau_{lt}) P_{ljt}^{1+\frac{\Lambda_{lt}}{1-\Lambda_{lt}}} - W_t P_{ljt}^{\frac{\Lambda_{lt}}{1-\Lambda_{lt}}} \right) P_{lt}^{-1-\frac{\Lambda_{lt}}{1-\Lambda_{lt}}} W_t \right] \\
&= \sum_{l\in[\mathscr{L}_j]} \zeta_l \theta_{lj} \left((1-\tau_{lt}) P_{ljt}^{1+\frac{\Lambda_{lt}}{1-\Lambda_{lt}}} - W_t P_{ljt}^{\frac{\Lambda_{lt}}{1-\Lambda_{lt}}} \right) P_{lt}^{-1-\frac{\Lambda_{lt}}{1-\Lambda_{lt}}}
\end{aligned}
$$

where I assume that the stochastic discount factor is given by $\sum_{l=1}^{n}\zeta_l \frac{U'(C_{lt})}{P_{lt}}$. Let

$$\pi_j((p_{ljt})_{l\in[n]};w_t,(p_{lt})_{l\in[n]},(\lambda_{lt})_{l\in[n]}) \equiv \Pi_j((P_{lj}e^{p_{ljt}})_{l\in[n]};We^{w_t},(P_l e^{p_{lt}})_{l\in[n]},(\Lambda_l e^{\lambda_{lt}})_{l\in[n]})$$
$$= \sum_{l\in[\mathscr{L}_j]}\zeta_l\theta_{lj}\Big((1-\tau_{lt})(P_{lj}e^{p_{ljt}})^{\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}} - We^{w_t}(P_{lj}e^{p_{ljt}})^{\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}}\Big)(P_l e^{p_{lt}})^{-1-\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}}$$

Now, define $\tilde{\vec{x}}_t = (w_t,p_{1t},p_{2t},\dots,p_{nt},\lambda_{1t},\lambda_{2t},\dots,\lambda_{nt})' \in \mathbb{R}^{2n+1}$ and $\vec{a}_{jt} = (p_{ljt})_{l\in[\mathscr{L}_j]}$. So

$$\pi^j(\vec{a}_{jt};\tilde{\vec{x}}_t) = \pi^j((p_{ljt})_{l\in[\mathscr{L}_j]};w_t,(p_{lt})_{l\in[n]},(\lambda_{lt})_{l\in[n]})$$

Let $\tilde{\vec{x}}^0 = \mathbf{0}_{(2n+1)\times 1}$. A second-order Taylor approximation of the profit function around $\tilde{\vec{x}}^0$ is

$$\pi^j(\vec{a}_{jt};\tilde{\vec{x}}_t) \approx \pi^j(\vec{a}^{\diamond}_{jt}(\tilde{\vec{x}}^0);\tilde{\vec{x}}^0) + \big(\boldsymbol{h}'_{\boldsymbol{a}}(\vec{a}^{\diamond}_{jt}(\tilde{\vec{x}}^0),\tilde{\vec{x}}^0),\boldsymbol{h}'_{\boldsymbol{x}}(\vec{a}^{\diamond}_{jt}(\tilde{\vec{x}}^0),\tilde{\vec{x}}^0)\big)\big(\vec{a}_{jt}-\vec{a}^{\diamond}_{jt}(\tilde{\vec{x}}^0),\tilde{\vec{x}}_{jt}-\tilde{\vec{x}}^0\big)$$
$$+ (\vec{a}'_{jt}-\vec{a}^{\diamond\prime}_{jt}(\tilde{\vec{x}}^0),\tilde{\vec{x}}'_{jt}-\tilde{\vec{x}}^{0\prime})\mathscr{H}_j\begin{pmatrix}\vec{a}_{jt}-\vec{a}^{\diamond}_{jt}(\tilde{\vec{x}}^0)\\ \tilde{\vec{x}}_{jt}-\tilde{\vec{x}}^0\end{pmatrix}$$
$$\approx \pi^j(\vec{a}^{\diamond}_{jt}(\tilde{\vec{x}}^0),\tilde{\vec{x}}^0) + \boldsymbol{h}'_{\boldsymbol{x}}(\vec{a}^{\diamond}_{jt}(\tilde{\vec{x}}^0),\tilde{\vec{x}}^0)(\tilde{\vec{x}}_t) + \frac{1}{2}(\vec{a}'_{jt})\mathscr{H}_{a,j}(\vec{a}_{jt}) + (\vec{a}'_{jt})\mathscr{H}_{ax,j}(\tilde{\vec{x}}_t) + \frac{1}{2}(\tilde{\vec{x}}'_t)\mathscr{H}_x(\tilde{\vec{x}}_t)$$

recalling that at the optimum $\boldsymbol{h}'_a(\vec{a}^{\diamond}_{jt}(\tilde{\vec{x}}^0),\tilde{\vec{x}}^0) = \vec{0}$, where

$$\mathscr{H}_j \equiv \begin{bmatrix}\mathscr{H}_{a,j} & \mathscr{H}_{ax,j}\\ \mathscr{H}_{xa,j} & \mathscr{H}_{x,j}\end{bmatrix}$$

with $\mathscr{H}_{a,j} = \big[\frac{\partial^2\pi}{\partial p_{ijt}\partial p_{kjt}}\big|_{\tilde{\vec{x}}^0}\big]_{i\in[\mathscr{L}_j],k\in[\mathscr{L}_j]}$ being the matrix of second derivatives of actions. $\mathscr{H}_{ax,j} = \big[\frac{\partial^2\pi}{\partial p_{ijt}\partial x_{kt}}\big]_{i\in[\mathscr{L}_j],k\in\{1,\dots,2n+1\}}$ is the matrix with cross derivatives of actions and states, and $\mathscr{H}_{x,j} = \big[\frac{\partial^2\pi}{\partial x_{kt}\partial x_{jt}}\big|_{\tilde{\vec{x}}^0}\big]_{k,j\in\{1,2,\dots,2n+1\}}$ the matrix of second derivatives of states. At the optimum, we have

$$\vec{a}^{\diamond}_{jt} = -\mathscr{H}^{-1}_{a,j}\mathscr{H}_{ax,j}\tilde{\vec{x}}_t$$

Now, $\vec{a}^*_{jt}$ be such $\vec{a}^*_{jt} = \text{argmax}_{\vec{a}}\mathbb{E}\big[\pi(\vec{a};\tilde{\vec{x}}_t)|\mathbf{S}^{jt}\big]$, $\mathbf{S}^{jt}$ being the history of signals up to time t. Then the profit losses arising from information frictions can be written as

$$\pi^j(\vec{a}^*_{jt};\tilde{\vec{x}}) - \pi^j(\vec{a}^{\diamond}_{jt};\tilde{\vec{x}}) = \frac{1}{2}\vec{a}^{*\prime}_{jt}\mathscr{H}_{a,j}\vec{a}^*_{jt} - \frac{1}{2}\vec{a}^{\diamond\prime}_{jt}\mathscr{H}_{a,j}\vec{a}^{\diamond}_{jt} + \vec{a}^{*\prime}_{jt}\mathscr{H}_{ax,j}\tilde{\vec{x}}_t - \vec{a}^{\diamond\prime}_{jt}\mathscr{H}_{ax,j}\tilde{\vec{x}}_t$$

Using $\vec{a}^{\diamond}_{jt} = -\mathscr{H}_{a,j}^{-1}\mathscr{H}_{ax,j}\tilde{\vec{x}}_t$, we have that $\vec{a}^{*}_{jt} = -\mathscr{H}_{a,j}^{-1}\mathscr{H}_{ax,j}\mathbb{E}[\tilde{\vec{x}}_t|\mathbf{S}^{jt}]$ and we can show that

$$\pi^j(\vec{a}^*_{jt};\tilde{\vec{x}}) - \pi^j(\vec{a}^{\diamond}_{jt};\tilde{\vec{x}}) = \frac{1}{2}(\vec{a}^{*\prime}_{jt} - \vec{a}^{\diamond\prime}_{jt})\mathscr{H}_{a,j}(\vec{a}^*_{jt} - \vec{a}^{\diamond}_{jt}) \tag{106}$$

$\mathscr{H}_{a,j}$ determines how mistakes in pricing for the many regions affect firm's profits. It also determines the extent to which mistakes in pricing in one of the regions are complements or substitutes to mistakes in pricing in other regions. Given the assumption of monopolistic competition, $\mathscr{H}_{a,j}$ is going to be diagonal, implying there is no complementarity or substitutability in actions. $-\mathscr{H}_{a,j}\mathscr{H}_{ax,j}$ the map of fundamental shocks into actions. Now, let's calculate these objects. Let $\pi^j_l(\vec{a}_{jt};\tilde{\vec{x}}_t) = \frac{\partial \pi^j}{\partial p_{ljt}}$, $\pi^j_{lm}(\vec{a}_{jt};\tilde{\vec{x}}_t) = \frac{\partial^2 \pi^j}{\partial p_{ljt}\partial p_{mjt}}$, $\pi^j_{lw}(\vec{a}_{jt};\tilde{\vec{x}}_t) = \frac{\partial^2 \pi^j}{\partial p_{ljt}\partial w_t}$, $\pi^j_{lp_m}(\vec{a}_{jt};\tilde{\vec{x}}_t) = \frac{\partial^2 \pi^j}{\partial p_{ljt}\partial p_{mt}}$, and $\pi^j_{l\lambda_m}(\vec{a}_{jt};\tilde{\vec{x}}_t) = \frac{\partial^2 \pi^j}{\partial p_{ljt}\partial \lambda_{mt}}$. Then,

$$\pi^j_l(\vec{a}_{jt};\tilde{\vec{x}}_t) = \zeta_l\theta_{lj}\Big((1-\tau_l)\Big(\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}\Big)(P_{lj}e^{p_{ljt}})^{\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}} - We^{w_t}\Big(\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}\Big)(P_{lj}e^{p_{ljt}})^{\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}}\Big)(P_l e^{p_{lt}})^{-1-\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}}$$

$$\pi^j_{lm}(\vec{a}_{jt};\tilde{\vec{x}}_t) =$$
$$\begin{cases} \zeta_l\theta_{lj}\Big(\Big(\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}\Big)^2(1-\tau_l)(P_{lj}e^{p_{ljt}})^{\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}} - We^{w_t}\Big(\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}\Big)^2(P_{lj}e^{p_{ljt}})^{\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}}\Big)(P_l e^{p_{lt}})^{-1-\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}} & \text{, if } l \in [\mathscr{L}_j], m = l \\ 0 & \text{, otherwise} \end{cases}$$

$$\pi^j_{lw}(\vec{a}_{jt};\tilde{\vec{x}}_t) = -\zeta_l\theta_{lj}\Big(We^{w_t}\Big(\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}\Big)(P_{lj}e^{p_{ljt}})^{\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}}\Big)(P_l e^{p_{lt}})^{-1-\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}}$$

$$\pi^j_{lp_m}(\vec{a}_{jt};\tilde{\vec{x}}_t) =$$
$$\begin{cases} \frac{1}{1-\Lambda_l e^{\lambda_{lt}}}\zeta_l\theta_{lj}\Big((1-\tau_l)\Big(\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}\Big)(P_{lj}e^{p_{ljt}})^{\frac{1}{1-\Lambda_l e^{\lambda_{lt}}}} - We^{w_t}\Big(\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}\Big)(P_{lj}e^{p_{ljt}})^{\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}}\Big)(P_l e^{p_{lt}})^{-1-\frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}} & \text{, if } l \in [\mathscr{L}_j], m = l \\ 0 & \text{, otherwise} \end{cases}$$

$$\pi^j_{l\lambda_m}(\vec{a}_{jt};\tilde{\vec{x}}_t) =$$
$$\begin{cases} \zeta_l\theta_{lj}\Big((1-\tau_l)P_{lj}e^{p_{ljt}}(P_l e^{p_{lt}})^{-1}e^{\mathscr{T}_{lt}}\Big[(g'(\lambda_{lt}))(P_{lj}e^{p_{ljt}}(P_l e^{p_{lt}})^{-1})^{g(\lambda_{lt})} + (1+g(\lambda_{lt}))(P_{lj}e^{p_{ljt}}(P_l e^{p_{lt}})^{-1})^{g(\lambda_{lt})}g'(\lambda_{lt})\ln(P_{lj}e^{p_{ljt}}(P_l e^{p_{lt}})^{-1})\Big] \\ \quad -We^{w_t}(P_l e^{p_{lt}})^{-1}\Big[g'(\lambda_{lt})(P_{lj}e^{p_{ljt}}(P_l e^{p_{lt}})^{-1})^{g(\lambda_{lt})} + g(\lambda_{lt})(P_{lj}e^{p_{ljt}}(P_l e^{p_{lt}})^{-1})^{g(\lambda_{lt})}g'(\lambda_{lt})\ln(P_{lj}e^{p_{ljt}}(P_l e^{p_{lt}})^{-1})\Big]\Big) & \text{, if } l \in [\mathscr{L}_j], m = l \\ 0 & \text{, otherwise} \end{cases}$$

where $g(\lambda_{lt}) \equiv \frac{\Lambda_l e^{\lambda_{lt}}}{1-\Lambda_l e^{\lambda_{lt}}}$. Now, evaluating these objects at $\tilde{\vec{x}}_t^0 = \mathbf{0}$, that is, around the non-stochastic steady state, we have

$$\pi^j_{lm}(\mathbf{0};\mathbf{0}) = \begin{cases} -\zeta_l \theta_{lj}\left(\frac{\Lambda_l}{\Lambda_l - 1}\right) & \text{, if } l \in [\mathscr{L}_j], m = l \\ 0 & \text{, otherwise} \end{cases}$$

and $\mathscr{H}_{a,j}$ is a diagonal matrix.

$$\pi^j_{lw}(\mathbf{0};\mathbf{0}) = \zeta_l \theta_{lj}\left(\frac{\Lambda_l}{\Lambda_l - 1}\right)$$

$$\pi^j_{lp_m}(\mathbf{0};\mathbf{0}) = 0, \forall m \in [n]$$

$$\pi^j_{l\lambda_m}(\mathbf{0};\mathbf{0}) = \begin{cases} \zeta_l \theta_{lj}\frac{\Lambda_l}{(\Lambda_l - 1)} & \text{, if } l \in [\mathscr{L}_j], m = l \\ 0 & \text{, otherwise} \end{cases}$$

Now, we can calculate $\vec{a}^{\diamond}_{jt} = (p^{\diamond}_{ljt})_{l\in[\mathscr{L}_j]}$ and $\vec{a}^{*}_{jt} = (p^{*}_{ljt})_{l\in[\mathscr{L}_j]}$

$$p^{\diamond}_{ljt} = w_t + \lambda_{lt}, \quad l \in [\mathscr{L}_j]$$

$$p^{*}_{ljt} = \mathbb{E}[w_t + \lambda_{lt} | \mathbf{S}^{jt}], \quad l \in [\mathscr{L}_j]$$

With the equations above and Equation (106) we are ready to solve the Rational Inattention problem.

Firm's Rational Inattention problem. The rational inattention firm's problem is given by

$$\max_{\{S_{j,t}\subset\mathbb{S}^t, \{p_{ljt}(S_j^t)\}_{l\in\mathscr{L}_j}\}_{t\geq 0}} \quad \mathbb{E}\Big[\sum_{t=0}^{\infty}\sum_{l\in\mathscr{L}_j} -\frac{B_{lj}}{2}(p_{ljt} - p^{\diamond}_{ljt})^2\Big] \tag{107}$$

$$\text{s.t.} \quad \mathbb{I}(S_j^t; \vec{x}^t | S_j^{t-1}) \leq \kappa_j \qquad \text{(info. processing constraint)} \tag{108}$$

$$S_j^t = S_j^{t-1} \cup S_{jt}, \quad S_j^{-1} \text{ given} \quad \text{(evolution of information set)} \tag{109}$$

with $B_{lj} = \zeta_l \theta_{lj}\frac{\Lambda_l}{\Lambda_l - 1}$

H Model Simulation

In this section, I describe the simulation of the panel of firms that is used to calibrate κ and to run the model validation regression. The quantitative model features ex-ante heterogeneous firms, where each different firm ex-ante heterogeneity is summarized by the set of of regions where it operates $\iota \in \{1,0\}^n \setminus \{\vec{0}\}$. First, for each different ex-ante heterogeneity ι with positive measure, $\varphi_\iota > 0$, I solve for the steady state information structure $(\bar{\Sigma}_{-1}(\iota), \bar{\Sigma}(\iota), \bar{\Omega}(\iota), \bar{\omega}(\iota))$ using DRIP from Afrouzi and Yang (2021), where $\bar{\omega}(\iota)$ is the steady-state lagrange multiplier associated with the capacity constraint of firm ex-ante heterogeneity ι. Then, I simulate N_f firms for $T + T_{\text{burn}}$ periods, where I burn the first T_{burn} periods. Let p_{ljt} be the price in region l, for firm j at time t. I end up with a balanced panel of firm-location prices, described by $\{p_{ljt}\}_{T_{\text{burn}+1} \leq t \leq T, j \in \{1,\ldots,N_f\}, l \in [\mathscr{L}_j]}$. At the end, we have $\varphi_\iota N_f$ firms with ex-ante heterogeneity ι, with $\sum_\iota \varphi_\iota N_f = N_f$. I choose $N_f = 5000$, $T = 500$, and $T_{\text{burn}} = 100$. To simulate $\{p_{ljt}\}_{0 \leq t \leq T}$ we need to simulate time series of the fundamental shocks $\{\vec{x}_t\}_{0 \leq t \leq T}$ and rational inattention shocks $\{\vec{v}_{jt}\}_{0 \leq t \leq T, j \in \{1,\ldots,N_f\}}$, where $\vec{x}_t = (m_t, \lambda_{1t}, \ldots, \lambda_{nt})'$, and $\vec{v}_{jt}$ is a vector with dimension that depends on the number of signals that the firm j chooses to receive. Note that under full information, $\{\vec{x}_t\}_{0 \leq t \leq T}$ would be sufficient to describe firms' prices. Before describing the simulation, let's define some objects.

Let $\mathbf{H}(\iota)$ be the map between states and actions for ι. That is, $\vec{p}^{\diamond}_{jt} = \mathbf{H}'(\iota)\vec{x}_t$, $j \in \iota$. Let the law of motion of fundamental shocks be given by

$$\vec{x}_t = \mathbf{A}\vec{x}_{t-1} + \mathbf{Q}\vec{u}_t, \quad \vec{u}_t \sim \mathcal{N}(\vec{0}, \mathbf{I})$$

For each ι, such that $\varphi_\iota > 0$, consider the set of optimal signals associated with the steady-state information structure $(\bar{\Sigma}_{-1}(\iota), \bar{\Sigma}(\iota), \bar{\Omega}(\iota), \bar{\omega}(\iota))$, $\{s_{ij,t}(\iota)\}_{1 \leq i \leq k(\iota)}$ for a firm j with type ι:

$$s_{ij,t}(\iota) = \mathbf{g}_i'(\iota)\vec{x}_t + v_{ij,t}(\iota), \quad v_{ij,t}(\iota) \sim \mathcal{N}(0, V(v_i(\iota)))$$

where $k(\iota)$ is the number of signals that type ι acquires, $g_i(\iota)$ is the vector of loadings associated with signal i for type ι, and $v_{ij,t}$ is the rational inattention error that firm j of type ι receives associated with signal i. Note that the only element that makes the signal to depend on j is the realization of $v_{ij,t}$. Finally, let d_i be the eigenvalue of the matrix $\bar{\Sigma}(\iota)^{\frac{1}{2}}_{-1}\bar{\Omega}(\iota)\bar{\Sigma}(\iota)^{\frac{1}{2}}_{-1}$ associated with the signal i. Now, we can describe the simulation.

1. Draw $\{\vec{u}_t\}_{1 \leq t \leq T+T_{\text{burn}}}$ from $\mathcal{N}(\vec{0}, \mathbf{I})$. Then, for $t \in \{1, \ldots, T + T_{\text{burn}}\}$, calculate $\vec{x}_t = \mathbf{A}\vec{x}_{t-1} + \mathbf{Q}\vec{u}_t$,

where $\vec{x}_{-1} = \vec{0}$.

2. Calculate $\{N_f(\iota)\}_\iota$, where $N_f(\iota) = \varphi_\iota N_f$ is the number of firms with type ι.

3. For each ι such that $\varphi_\iota > 0$,

   (a) Draw time series of Rational Inattention shocks for firms of type ι, $\{\vec{\nu}_{jt}(\iota)\}_{j\in\{1,\dots,N_f(\iota)\},1\leq t\leq T+T_{\text{burn}}}$

   (b) For each firm $j \in \{1,\dots,N_f(\iota)\}$,

      i. Calculate the law of motion of conditional expectations

$$\mathbb{E}[\vec{x}_t|\mathbf{S}_j^t(\iota)] = \mathbf{A}\mathbb{E}[\vec{x}_{t-1}|\mathbf{S}_j^{t-1}(\iota)] + \sum_{i=1}^{k(\iota)}\left(1-\frac{\bar{\omega}(\iota)}{d_i(\iota)}\right)\bar{\Sigma}_{-1}(\iota)\mathbf{g}_i(\iota)(\mathbf{g}_i'(\iota)(\vec{x}_t - \mathbf{A}\mathbb{E}[\vec{x}_{t-1}|\mathbf{S}_j^{t-1}]) + \nu_{ij,t}(\iota))$$

      where $\mathbf{S}_j^t(\iota) = \mathbf{S}_j^{t-1}(\iota)\cup\{s_{ijt}(\iota)\}_{1\leq i\leq k(\iota)}$ is the history of signals up to t, and $\mathbb{E}[\vec{x}_{-1}|\mathbf{S}_j^{-1}(\iota)] = \vec{0}$

      ii. Calculate the optimal actions

$$\vec{p}_{jt}(\iota) = \mathbf{H}'(\iota)\mathbb{E}[\vec{x}_t|\mathbf{S}_j^t(\iota)]$$

I Data

In this section, I describe the details of the data used in Section 3.3 and Section 3.5.

I.1 The NielsenIQ Retail Scanner Data

The main dataset used in this paper is the NielsenIQ Retail Scanner Data (NielsenIQ). This dataset contains weekly scanner prices and quantities for different products (UPCs) for retail stores that share data with NielsenIQ for the whole U.S. market. I use data from 2006 to 2019 for food stores in the United States.

I restrict the set of chains and stores used in the analysis. I define a chain to be a unique combination of two identifiers in the NielsenIQ data: *parent_code* and *retailer_code*. The *parent_code* can be either the corporate parent or the retail banner of the store, depending on how the retailer releases the data to NielsenIQ. The *retailer_code* may or may not be the same as the *parent_code*. If a retailer has many banners, the *retailer_code* may be one of the banners. Over time, it may be the case that a given *retailer_code* is associated with more than one *parent_code*, which can occur because of ownership changes. A general example taken from DellaVigna and Gentzkow (2019) would be the Albertson's LLC parent company, which owns chains like Albertson's and Shaw's.

The stores used in my main analysis satisfy the following criteria: 1) I consider only food stores[15]; 2) I exclude stores without *retailer_code*; 3) Keep only stores with unique chains, that is keep stores with unique (parent_code, retailer_code) pairs; 4) I drop stores with less than or equal to two years in the data; 5) keep stores that are present in the data for at least 10 years. After applying these restriction, I am left with 9479 stores from 65 chains. We further restrict the sample of stores: 6) I exclude stores that change fips_code, defined as the combination of (fips_state_code,fips_county_code); 7) I keep only stores present in a core-based statistical area (CBSA)[16]. This leaves me with a final sample of 8859 stores from 65 chains.

I restrict the set of potential products used in the analysis. Each product is characterized by its UPC, which is a 12 numeric digit code associated with a barcode. The NielsenIQ Retail Scanner data classifies products in 10 different departments. Within each department, products are classified into different product groups, which are further classified into different modules. An example of a module, its product group, and its department is the module 'cereal - ready to eat', that belongs to the 'cereal' product group, that belongs to the 'dry grocery' department. I restrict the set of potential products I use in my analysis to those belonging to the following departments: dry grocery, frozen foods, dairy, deli, packaged meat, fresh produce, and alcoholic beverages [17]. In additional to these products, I include all product categories that are used in DellaVigna and Gentzkow (2019), but are not in these departments[18].

To assess how stringent were the data filters I applied, I construct a U.S. NielsenIQ price index following the methodology of Beraja, Hurst, and Ospina (2019). I then compare its year-over-year changes with the CPI for food at home, as illustrated in Figure J.23. The close alignment between the two series indicates that the sample I am using is representative of the prices included in the CPI for food at home.

I.2 Calibration of σ_u

To calibrate σ_u, I first download the U.S. nominal GDP series from St. Louis FRED (code: NGDP-SAXDCUSQ) for the period from January, 1990 until January, 2020. Then, I use a spline routine to

[15]The other possible channels are: convenience stores, drug stores, mass mechandisers, and liquor stores.

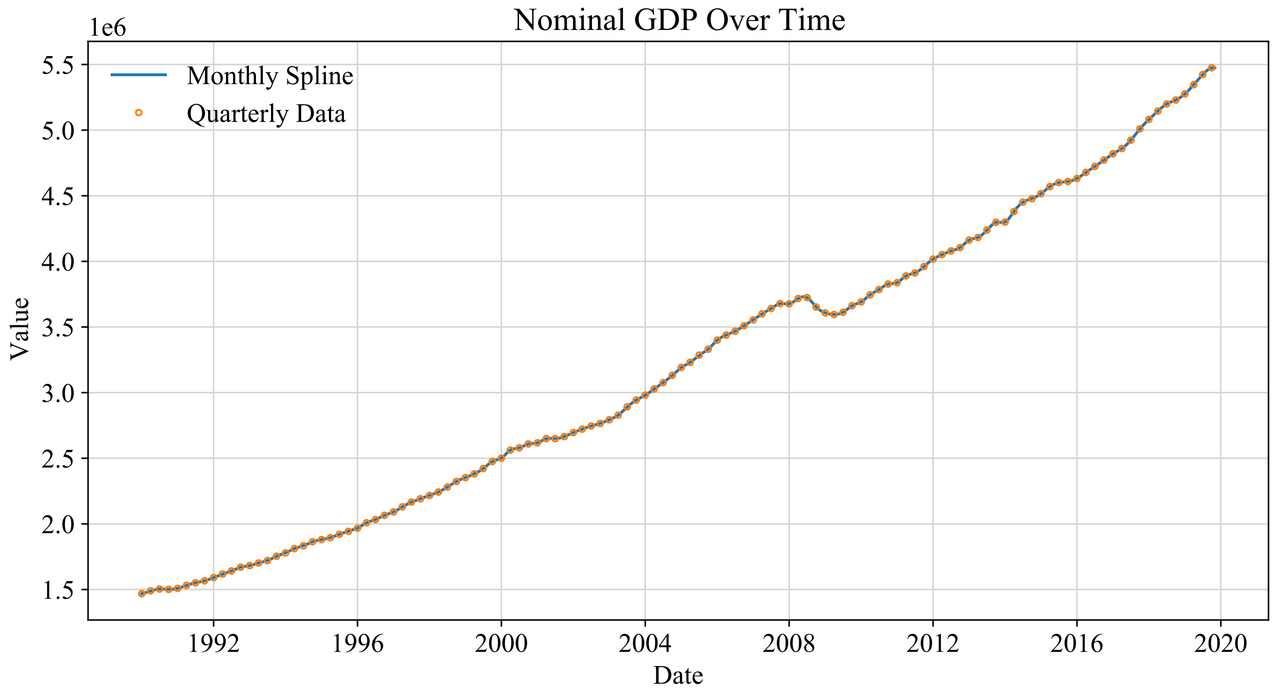
[16]I use CBSA codes from 2018.

[17]The other departments are: health & beauty care, non-food grocery, and general merchandise.

[18]The modules are toilet tissue, detergents - heavy duty - liquid, paper towels, cold remedies - adult, pain remedies - headache, batteries, bleach - liquid/gel

interpolate the nominal GDP series to get a monthly series. Then, a calculate the log difference and take its standard deviation. Figure I.1 shows both the raw data and the monthly spline series.

Figure I.1: Nominal GDP



Notes: This figure plots U.S. nominal GDP series (St. Louis FRED code: NGDPSAXDCUSQ) and its monthly interpolated series.

I.3 Calibration of $\{\sigma_l\}_{l\in\{1,...,12\}}$

To calibrate $\{\sigma_l\}_{l\in\{1,...,12\}}$ I use the Quarterly Census of Employment and Wages data from the Bureau of Labor Statistics for the period from January, 1990 until December, 2019. I consider the county level employment for private sector in all industries and aggregate it at the Federal Reserve district level, using the table mapping counties to Federal Reserve districts provided in https://www.kansascityfed.org/research/technical-briefings-sub/federal-reserve-district-county-shapefiles/. Then, seasonally adjust each Federal Reserve district employment level using X13.

Table I.1: Federal Reserve Districts Shock Volatilities

Fed District	Standard Deviation
Boston	0.002657
New York	0.002288
Philadelphia	0.002245
Cleveland	0.002306
Richmond	0.002535
Atlanta	0.002706
Chicago	0.004916
St. Louis	0.009098
Minneapolis	0.002602
Kansas City	0.005456
Dallas	0.002666
San Francisco	0.002716

I.4 Calibration of Taste Shifters $\{\theta_{lj}\}_{l\in\{1,...,12\},j\in[0,1]}$

In the efficient steady state, the household's expenditure share is given by

$$\frac{\theta_{lj}}{\int_0^1 \theta_{lj}dj} = \frac{P_{lj}C_{lj}}{P_lC_l}$$

where $\frac{P_{lj}C_{lj}}{P_lC_l}$ is the consumption share of location l on retailer j. I assume that $\int_0^1 \theta_{lj}dj = 1$, $\forall l \in [n]$. Furthermore, I assume that for each $l \in [n]$, firms that are present in the *same number* of locations have the same θ_{lj}. Therefore,

$$\theta_{l,k\text{ regions}} \times \int_0^1 \mathbf{1}\{j \text{ in } l \text{ and } k-1 \text{ other regions}\}dj = \frac{\int_0^1 P_{lj}C_{lj}\mathbf{1}\{j \text{ in } l \text{ and } k-1 \text{ other regions}\}dj}{P_lC_l}$$

$$\theta_{l,k\text{ regions}} \times \sum_{\iota\in l}\varphi_\iota = \frac{\text{Expenditure on chains in } l \text{ and } k-1 \text{ other regions}}{\text{Expenditure in } l}$$

where $\sum_{\iota\in l}\varphi_\iota$ is the measure of types that are present in l, $\mathbf{1}\{j \text{ in } l \text{ and } k-1 \text{ other regions}\}$ is a dummy that takes the value of one if j in l and $k-1$ other regions and zero otherwise.

I.5 Data Moment

The data moment I use to calibrate $m(\kappa)$ is given by

$$m_{\text{data}} = \sum_t \Big(\sum_j \Big[\sum_c \frac{var_{cjt}(\hat{p}_{cljt})}{var_{ct}(\hat{p}_{cljt})} \Big] \Big)$$

where g index a UPC, c a product category, l a county, j a chain, t a time period (month), $var_{cjt} \equiv \frac{1}{N_l - 1} \sum_l (\hat{p}_{cljt} - \frac{1}{N_l} \sum_l \hat{p}_{cljt})^2$, and $var_{ct} \equiv \frac{1}{N_{lj}-1} \sum_j \sum_l (\hat{p}_{cljt} - \frac{1}{N_{lj}} \sum_j \sum_l \hat{p}_{cljt})^2$, with N_l being the number of locations N_{lj} the number of locations and firms, with $\hat{p}_{cljt}$ is defined as

$$\hat{p}_{cljt} = \frac{1}{N_g} \sum_{g \in c} \hat{p}_{cljt}(g)$$

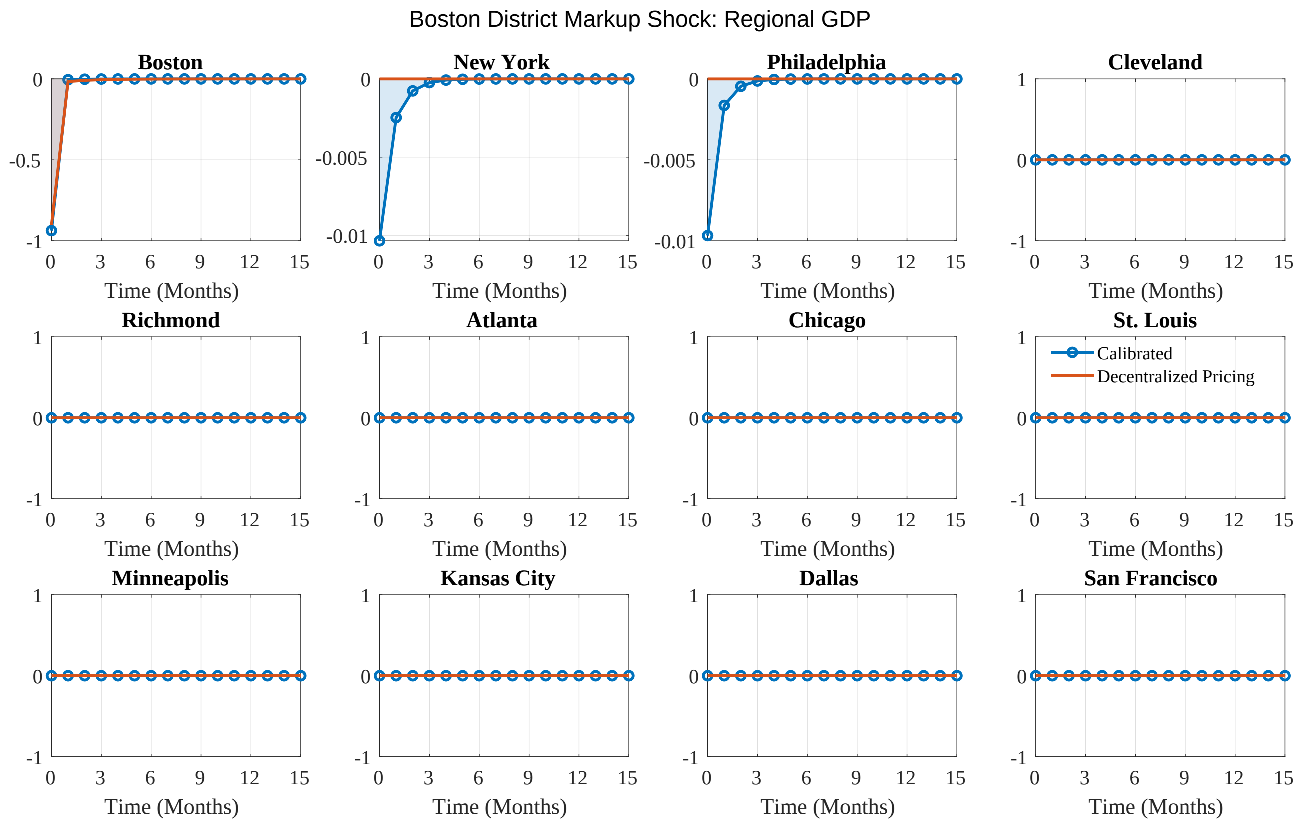
where N_g is the number of UPCs in category c for firm j in location l at time t, where I omit the fact that it depends on ljt. $\hat{p}_{cljt}(g)$ is given by

$$\hat{p}_{cljt}(g) = p_{cljt}(g) - \mu_g - \epsilon_{gt} - \mu_{cl}$$

where $\mu_g = \frac{1}{N_{tjl}} \sum_t \sum_j \sum_l p_{cljt}(g)$, $\epsilon_{gt} = \frac{1}{N_{jl}} \sum_j \sum_l (p_{cljt}(g) - \mu_g)$, $\mu_{cl} = \frac{1}{N_{tj}} \sum_t \sum_j (p_{cljt}(g) - \mu_g - \epsilon_{gt})$ and I calculate these objects within a category. N_{tjl} is the number periods, firms, and locations, recalling that firms may be present in different number of locations, for different number of periods, and good prices availability may vary along all these dimensions. Finally, $p_{cljt}(g)$ is the reference price for g in location l, firm j, and time period (month) t. That is, as in Eichenbaum, Jaimovich, and Rebelo (2011), for a good g, it is the most often quoted (mode) price for chain j, county l across all of its stores at month t. Therefore, we are collapsing store-week price level into chain-county-month level data. I only use observations with only one mode.

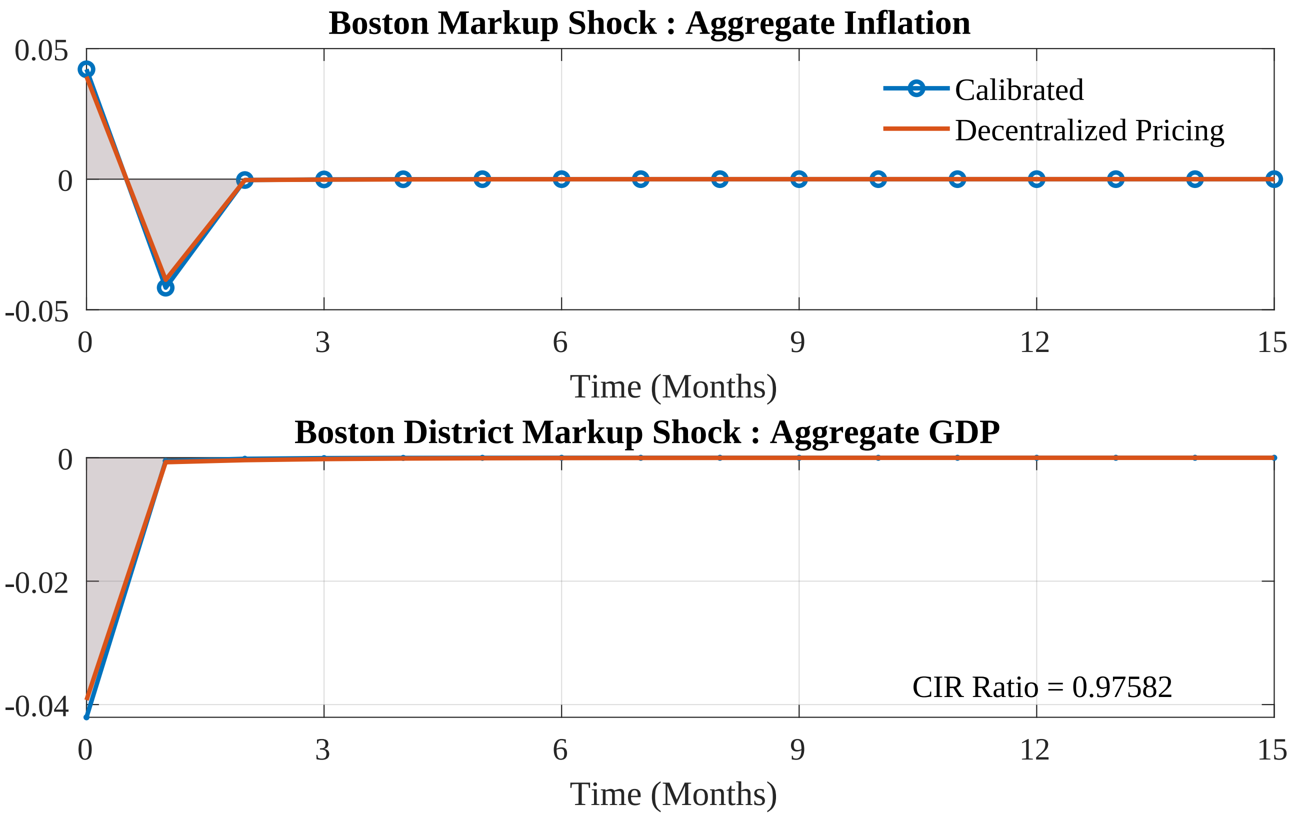
J Additional Figures

Figure J.1: Impulse response functions to a markup shock in the Boston Fed district



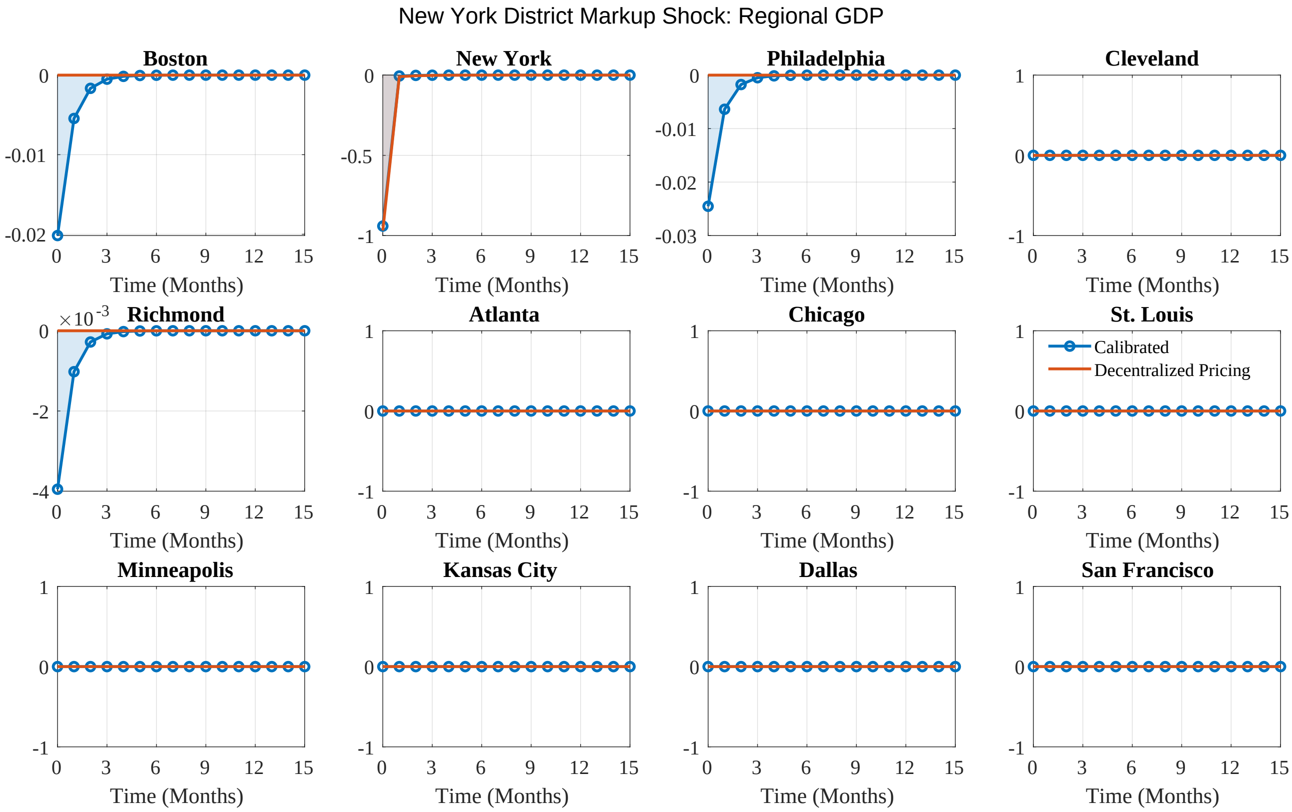
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Boston Fed district.

Figure J.2: Impulse response functions to a markup shock in the Boston Fed district



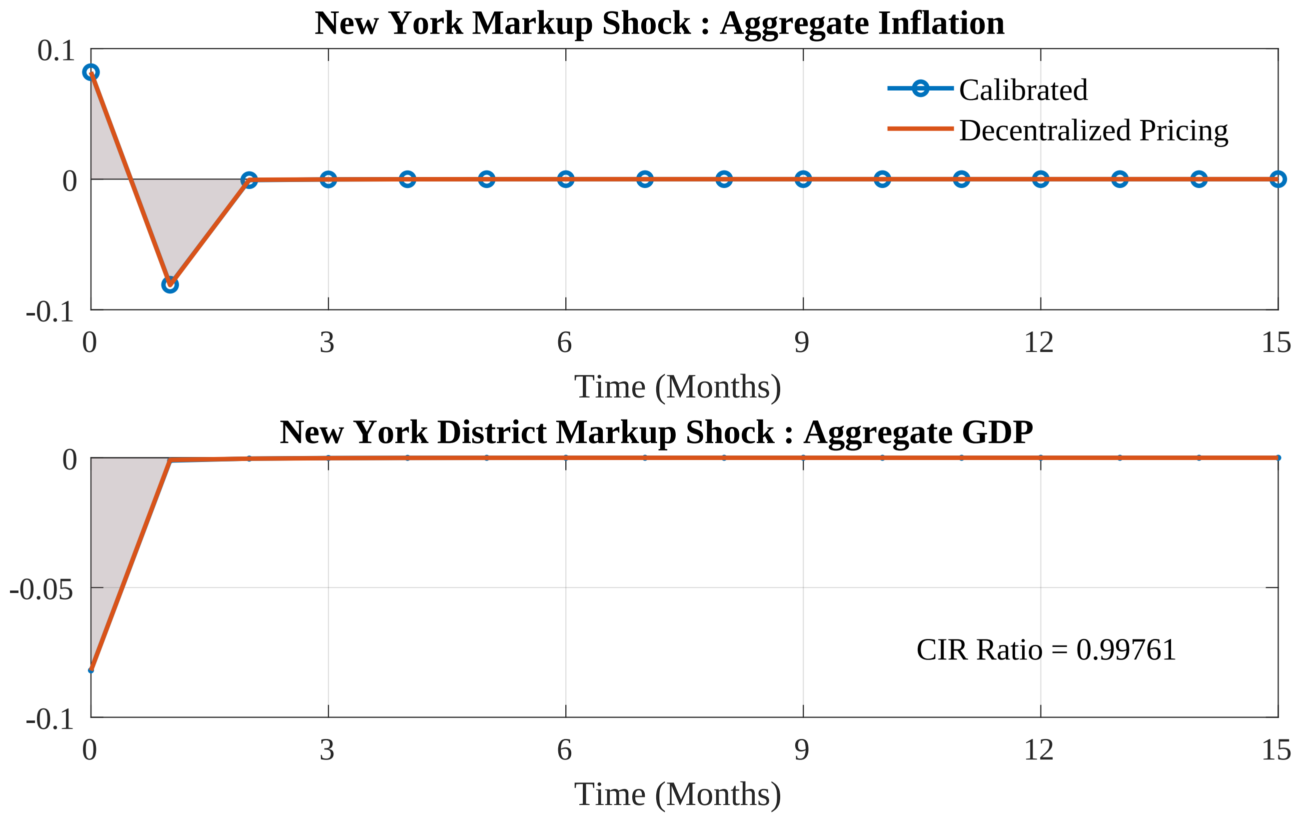
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Boston Fed district.

Figure J.3: Impulse response functions to a markup shock in the New York Fed district



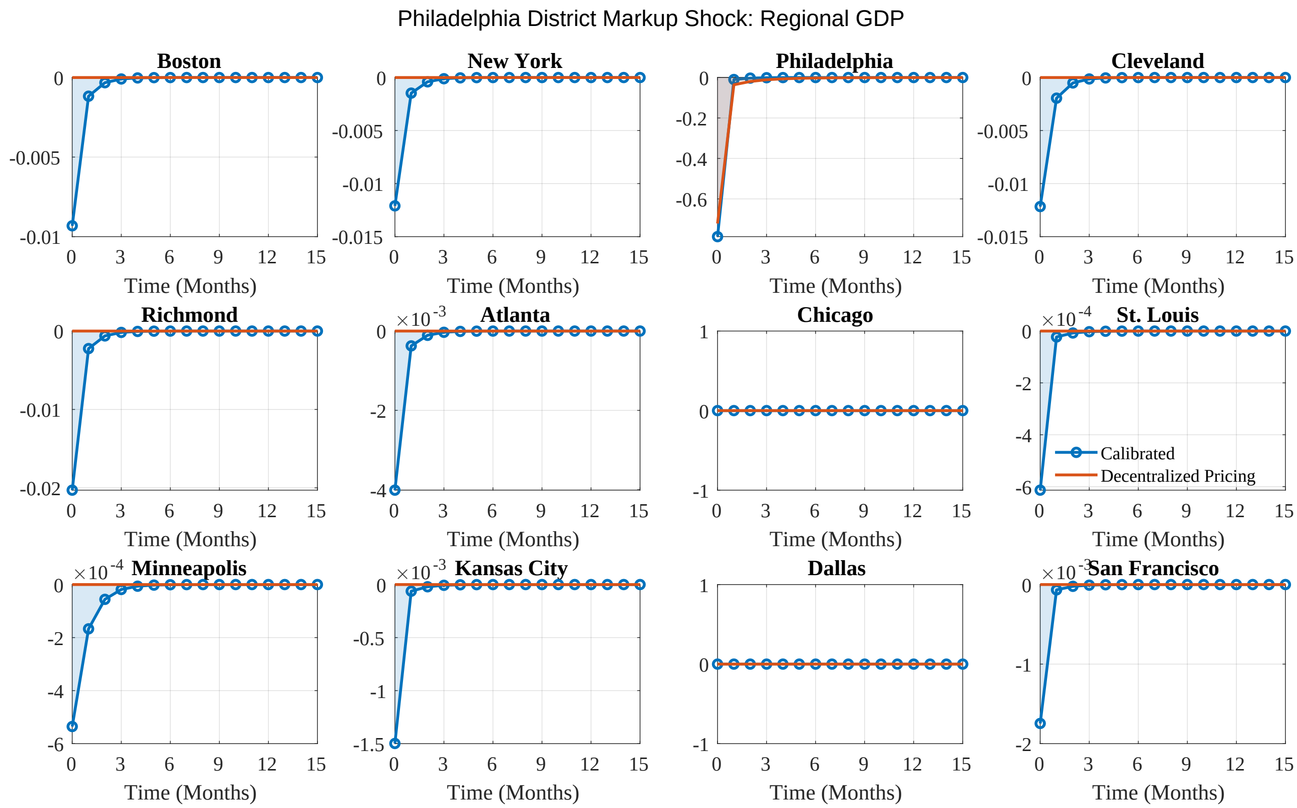
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. New York Fed district.

Figure J.4: Impulse response functions to a markup shock in the New York Fed district



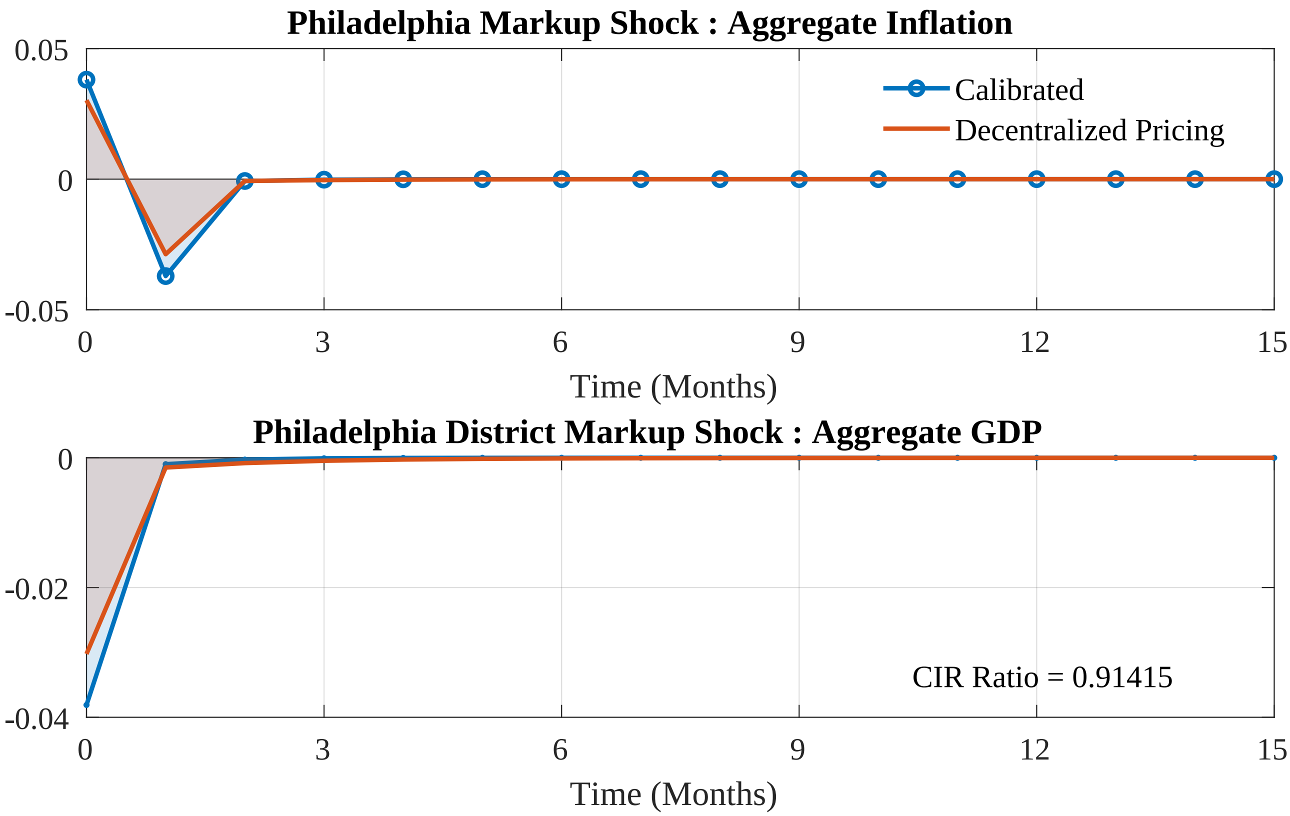
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. New York Fed district.

Figure J.5: Impulse response functions to a markup shock in the Philadelphia Fed district



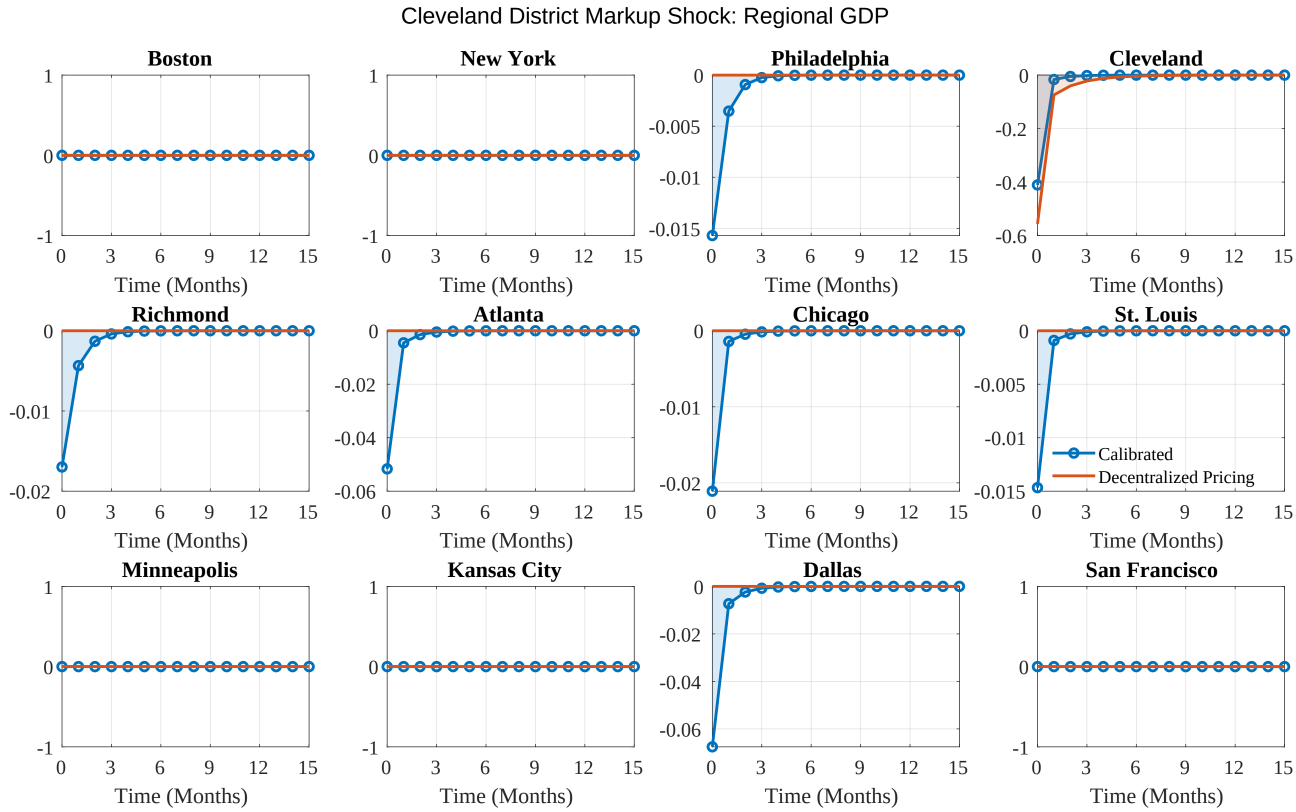
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Philadelphia Fed district.

Figure J.6: Impulse response functions to a markup shock in the Philadelphia Fed district



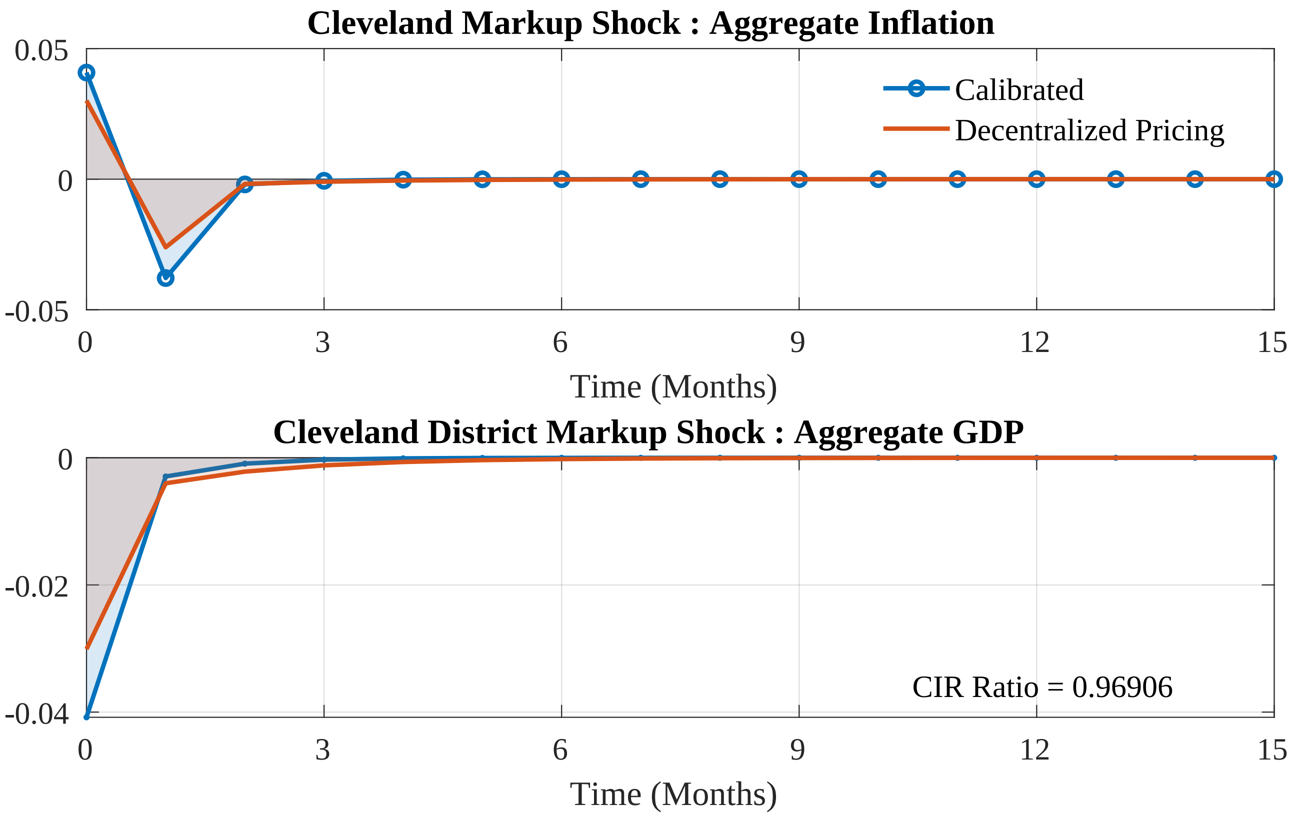
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Philadelphia Fed district.

Figure J.7: Impulse response functions to a markup shock in the Cleveland Fed district



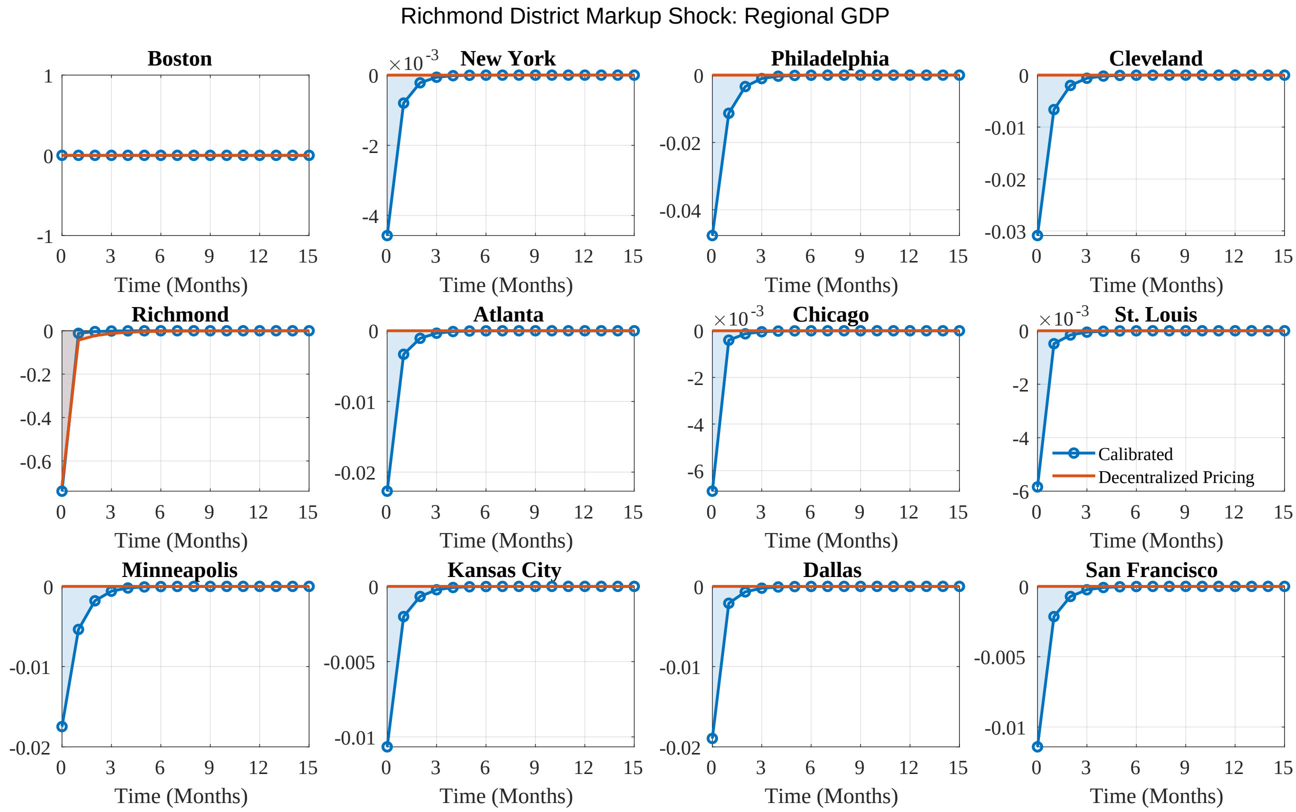
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Cleveland Fed district.

Figure J.8: Impulse response functions to a markup shock in the Cleveland Fed district



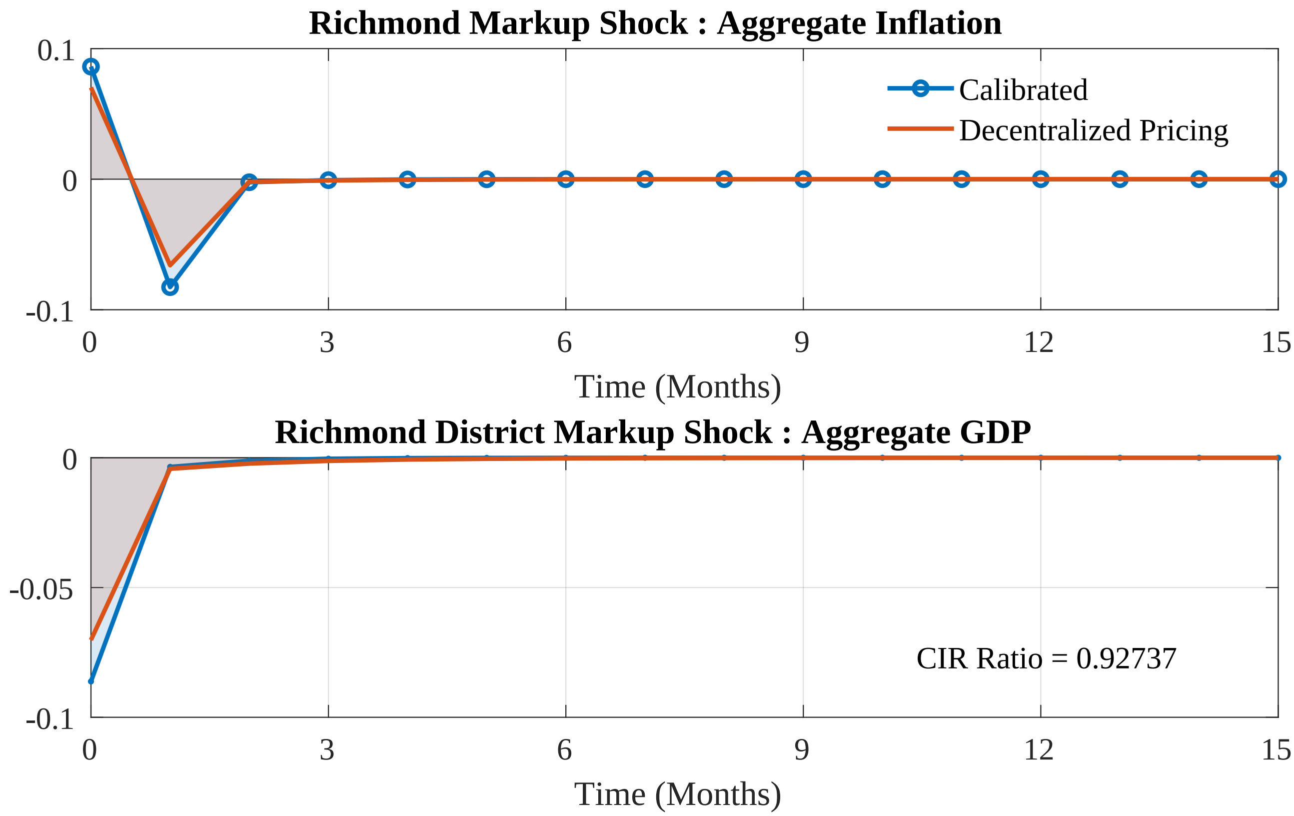
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Cleveland Fed district.

Figure J.9: Impulse response functions to a markup shock in the Richmond Fed district



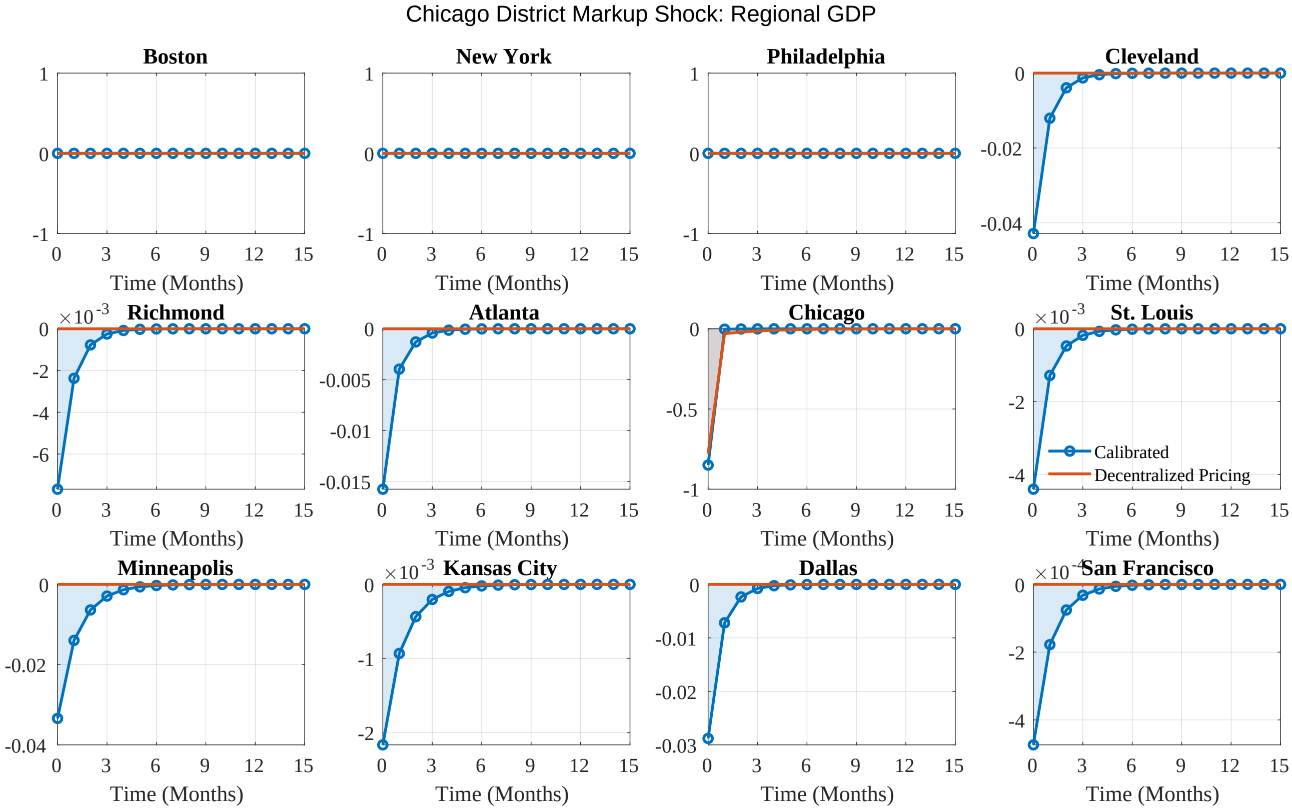
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Richmond Fed district.

Figure J.10: Impulse response functions to a markup shock in the Richmond Fed district



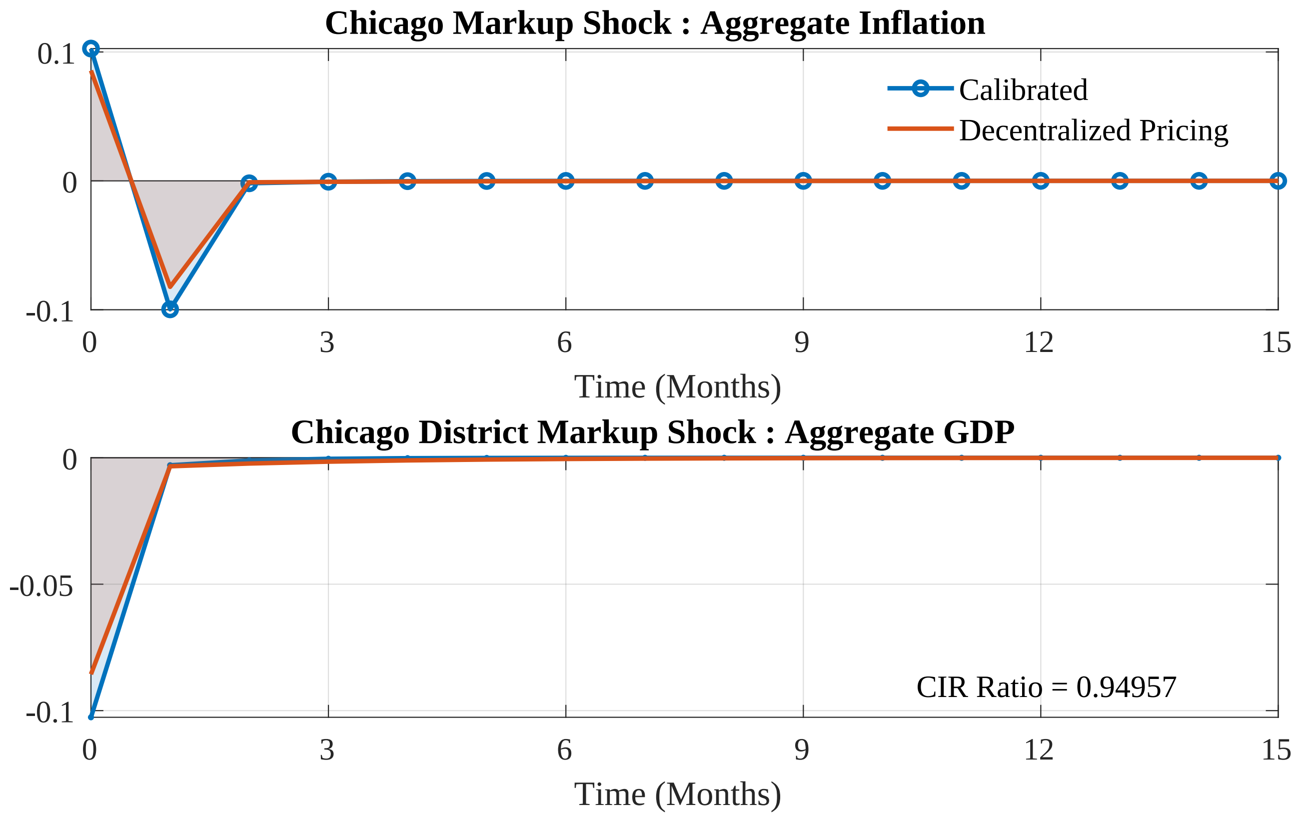
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Richmond Fed district.

Figure J.11: Impulse response functions to a markup shock in the Chicago Fed district



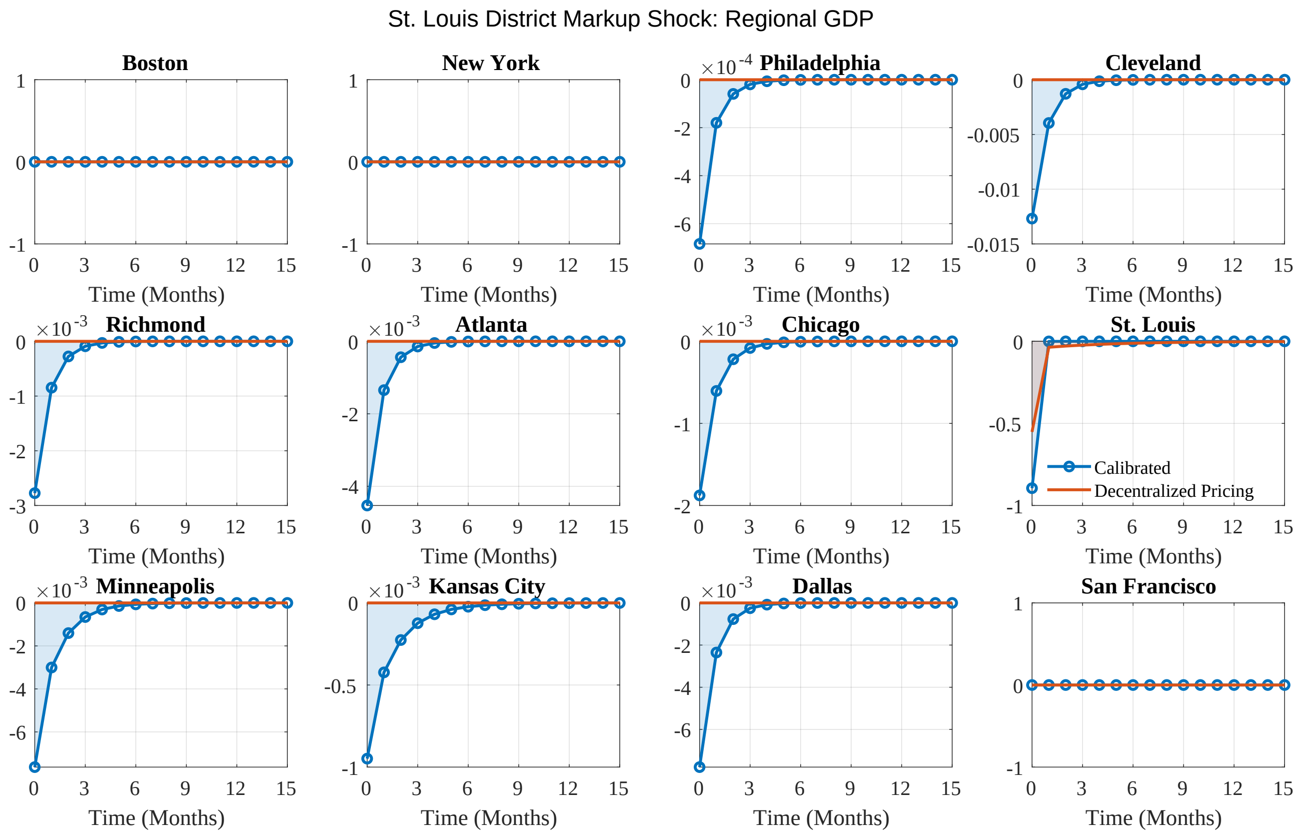
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Chicago Fed district.

Figure J.12: Impulse response functions to a markup shock in the Chicago Fed district



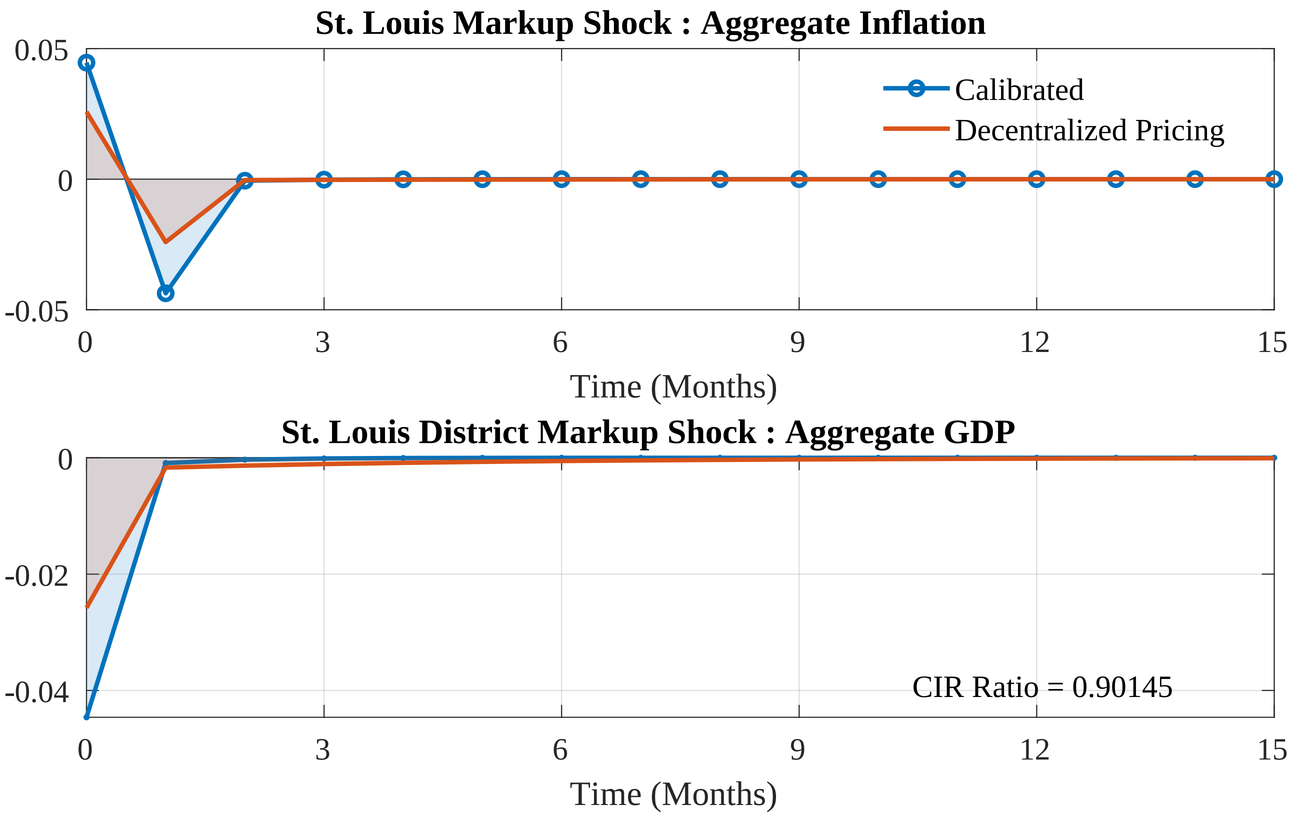
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Chicago Fed district.

Figure J.13: Impulse response functions to a markup shock in the St. Louis Fed district



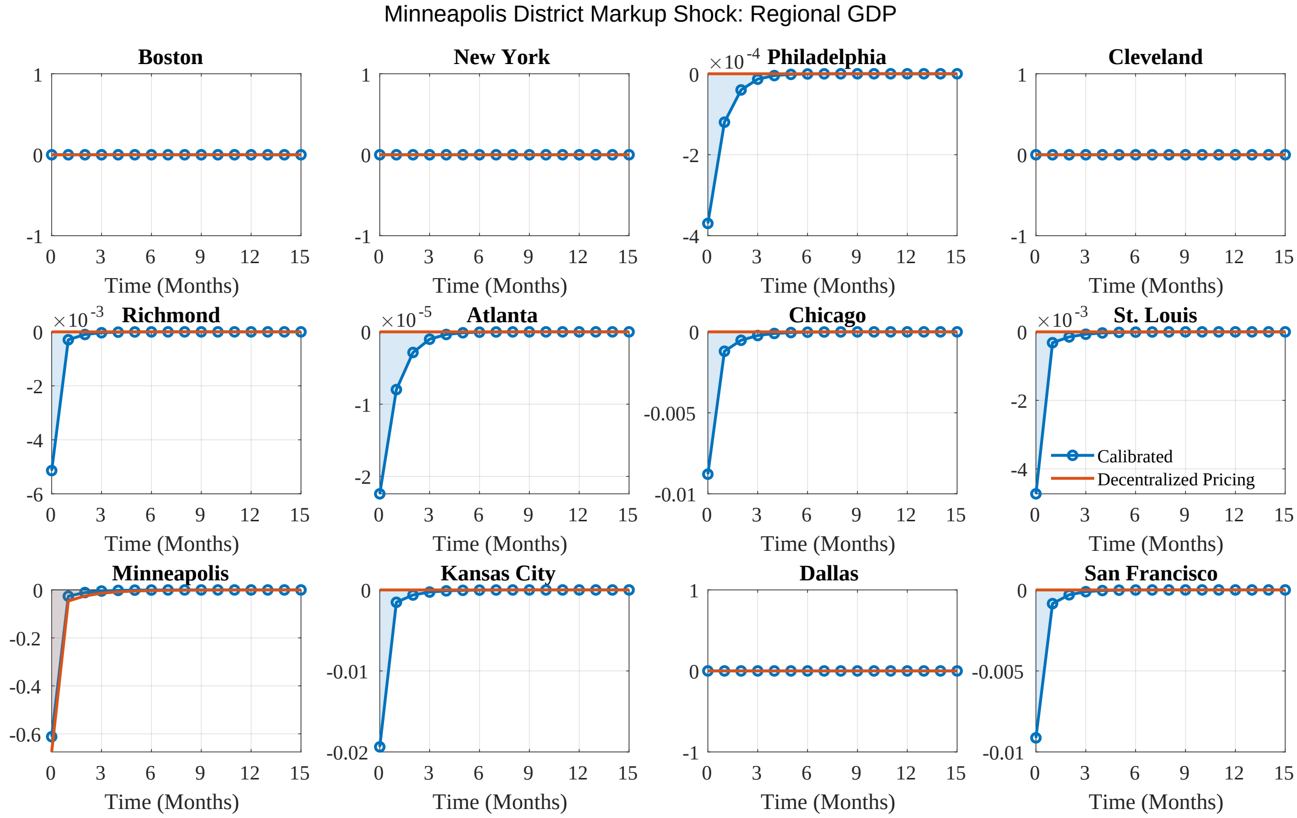
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. St. Louis Fed district.

Figure J.14: Impulse response functions to a markup shock in the St. Louis Fed district



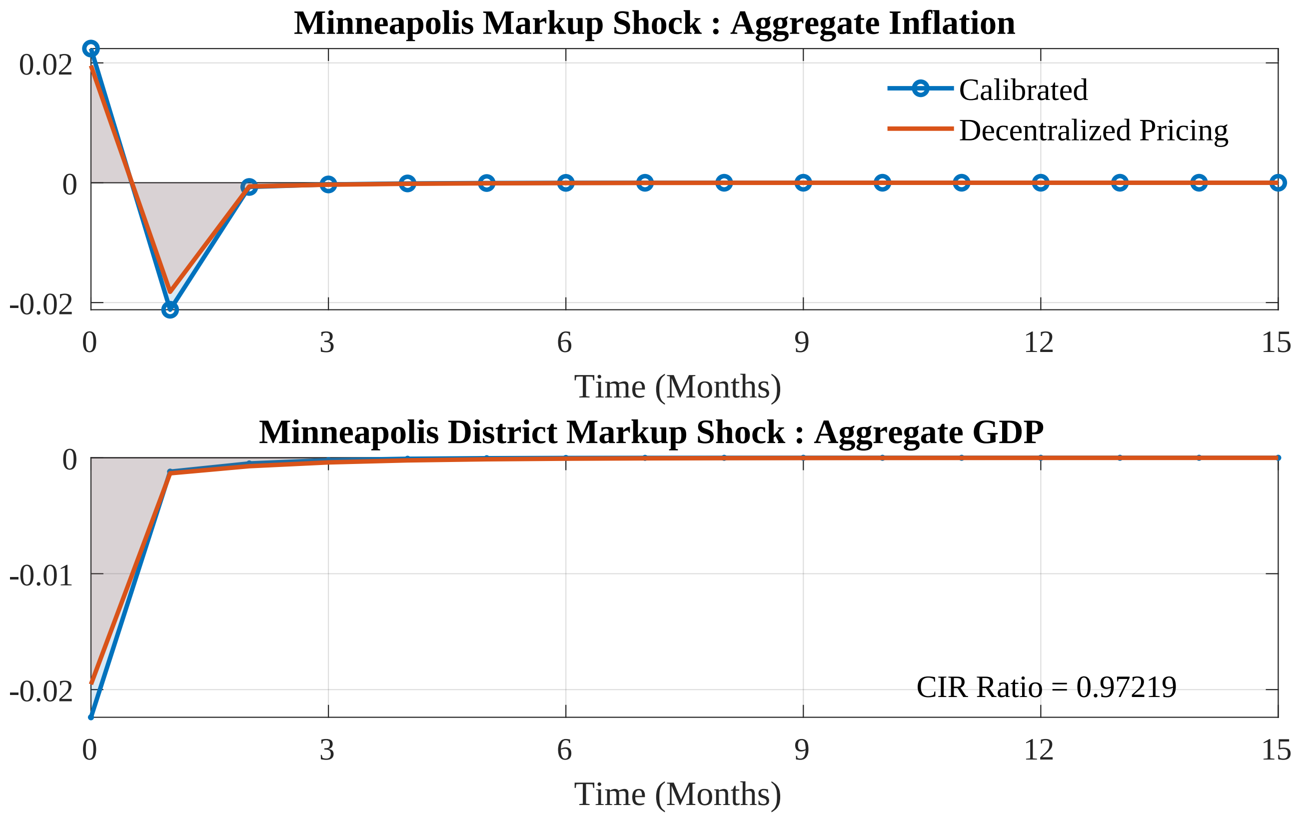
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. St. Louis Fed district.

Figure J.15: Impulse response functions to a markup shock in the Minneapolis Fed district



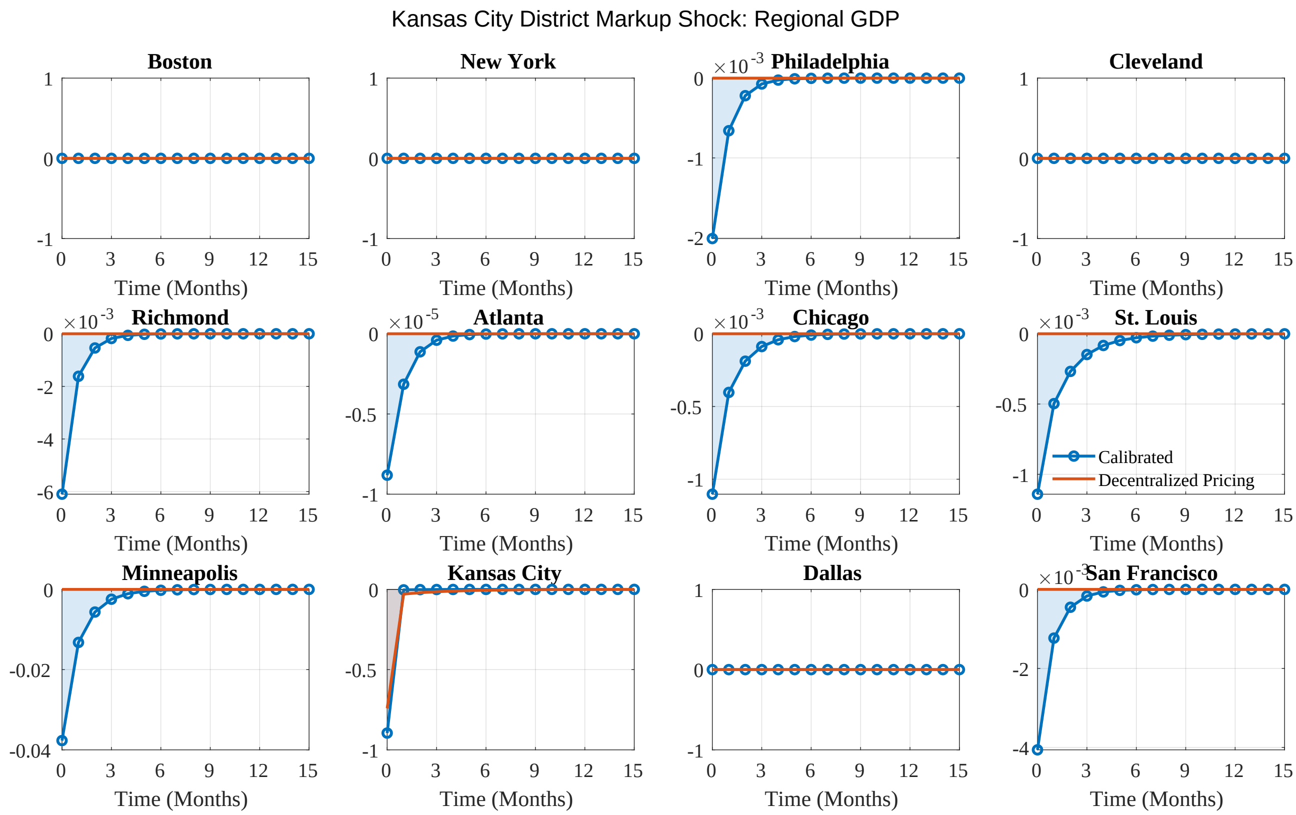
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Minneapolis Fed district.

Figure J.16: Impulse response functions to a markup shock in the Minneapolis Fed district



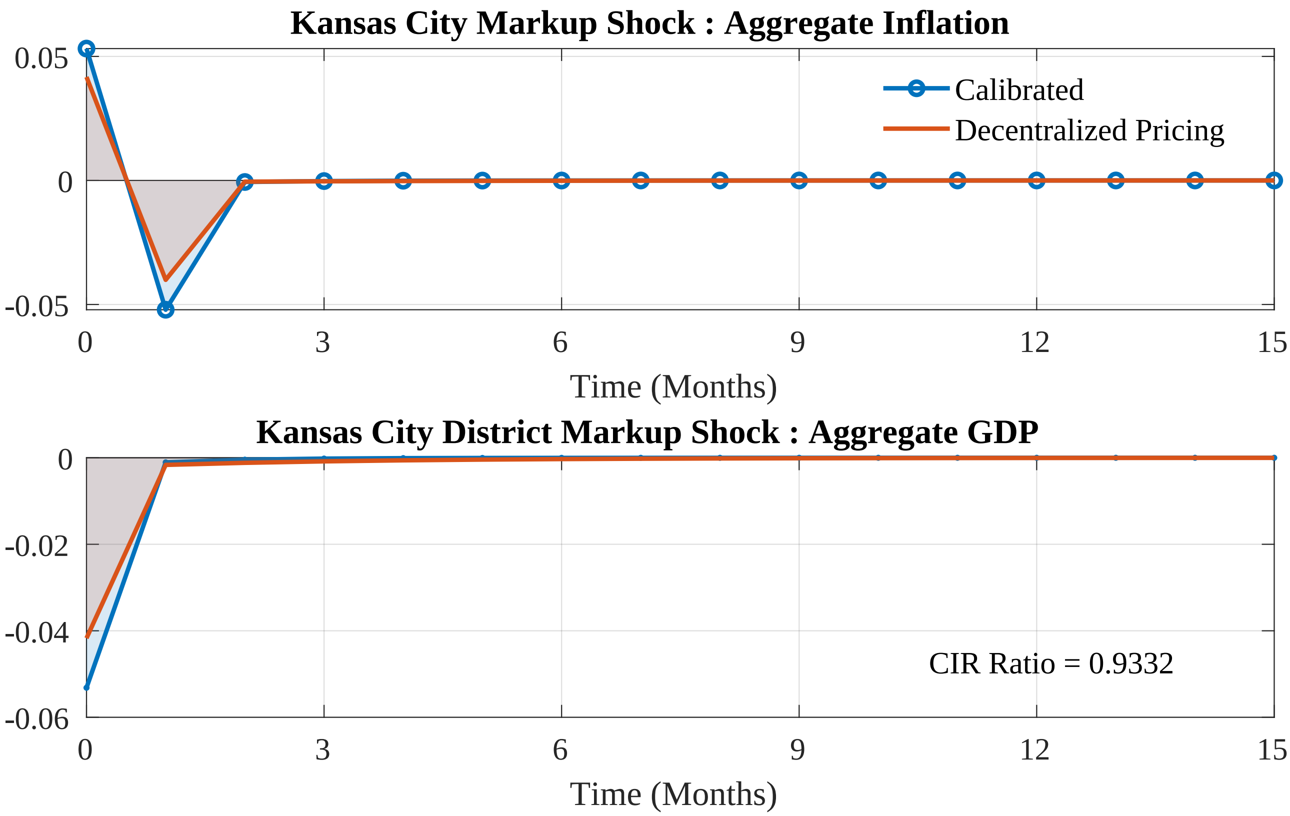
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Minneapolis Fed district.

Figure J.17: Impulse response functions to a markup shock in the Kansas City Fed district



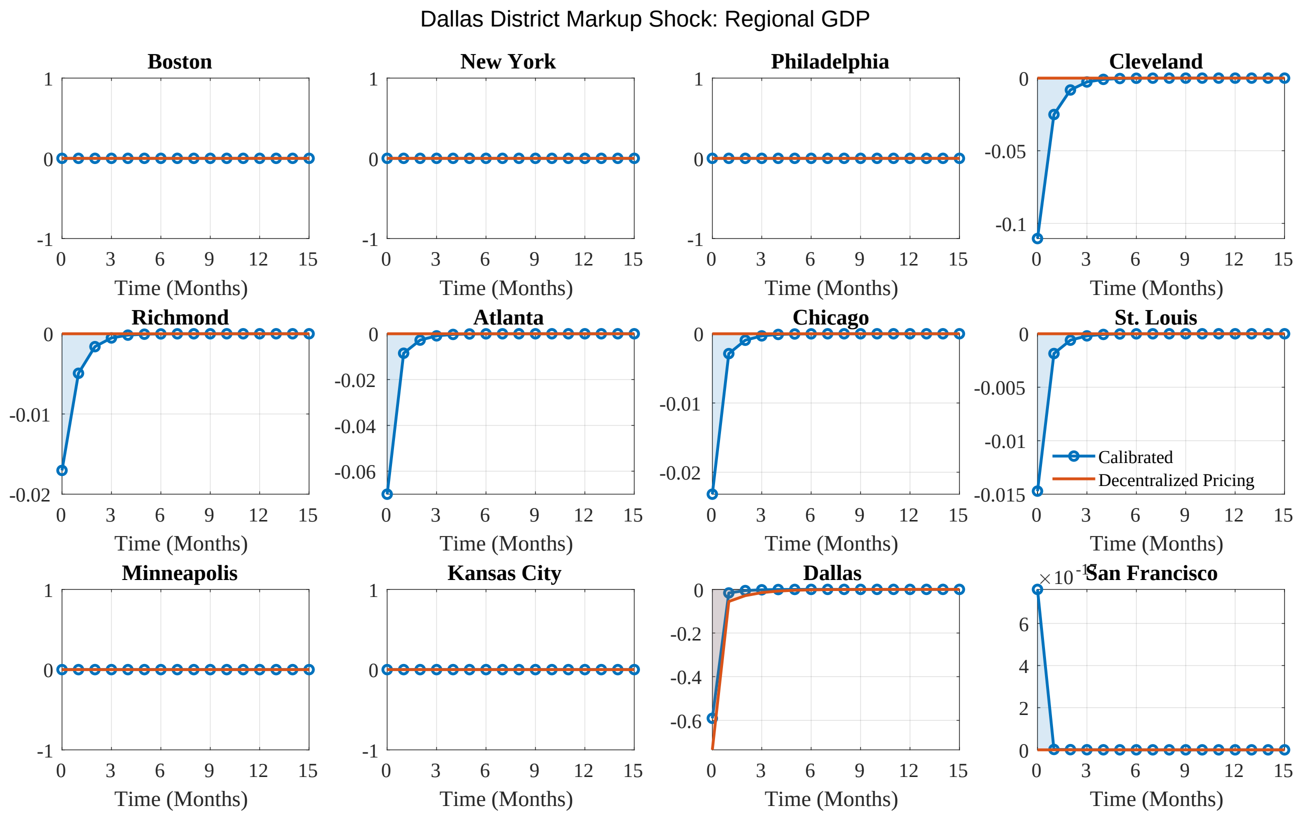
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Kansas City Fed district.

Figure J.18: Impulse response functions to a markup shock in the Kansas City Fed district



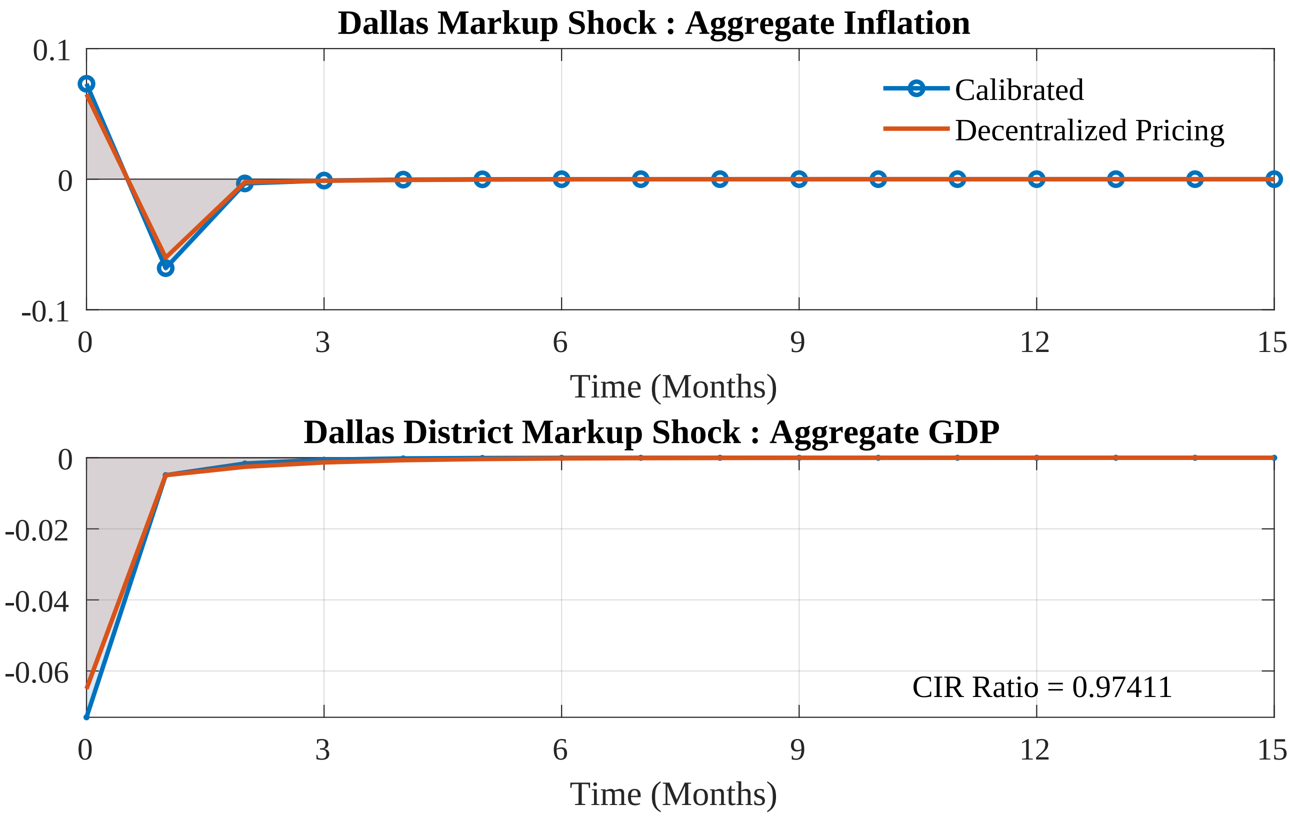
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Kansas City Fed district.

Figure J.19: Impulse response functions to a markup shock in the Dallas Fed district



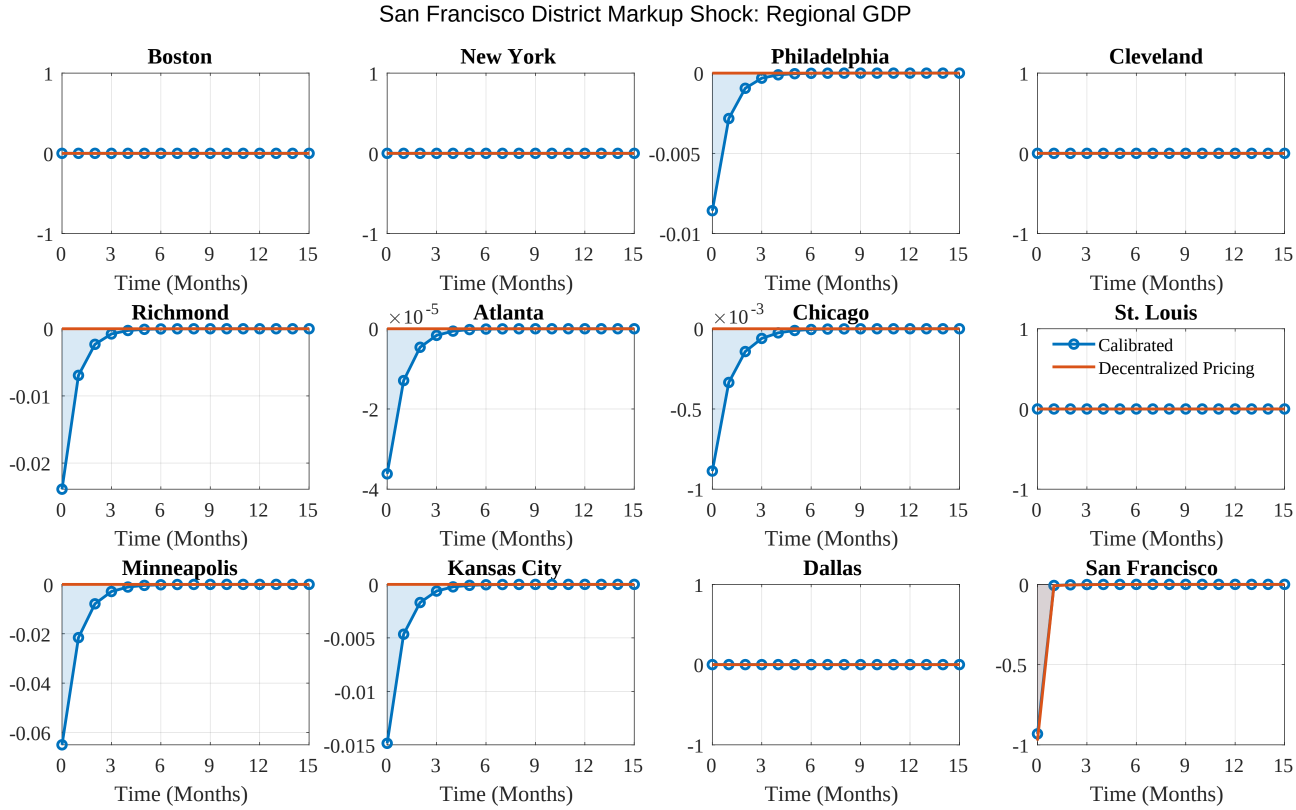
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Dallas Fed district.

Figure J.20: Impulse response functions to a markup shock in the Dallas Fed district



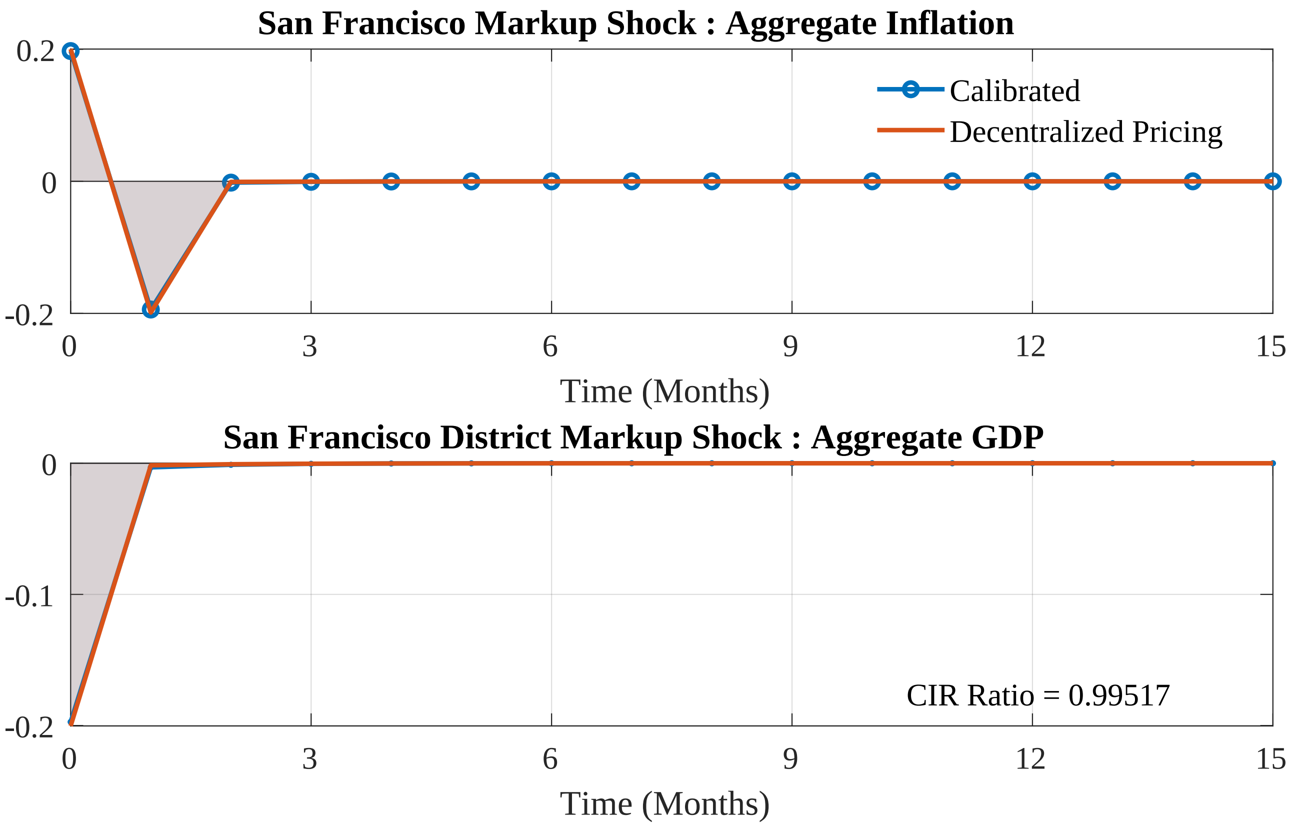
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. Dallas Fed district.

Figure J.21: Impulse response functions to a markup shock in the San Francisco Fed district



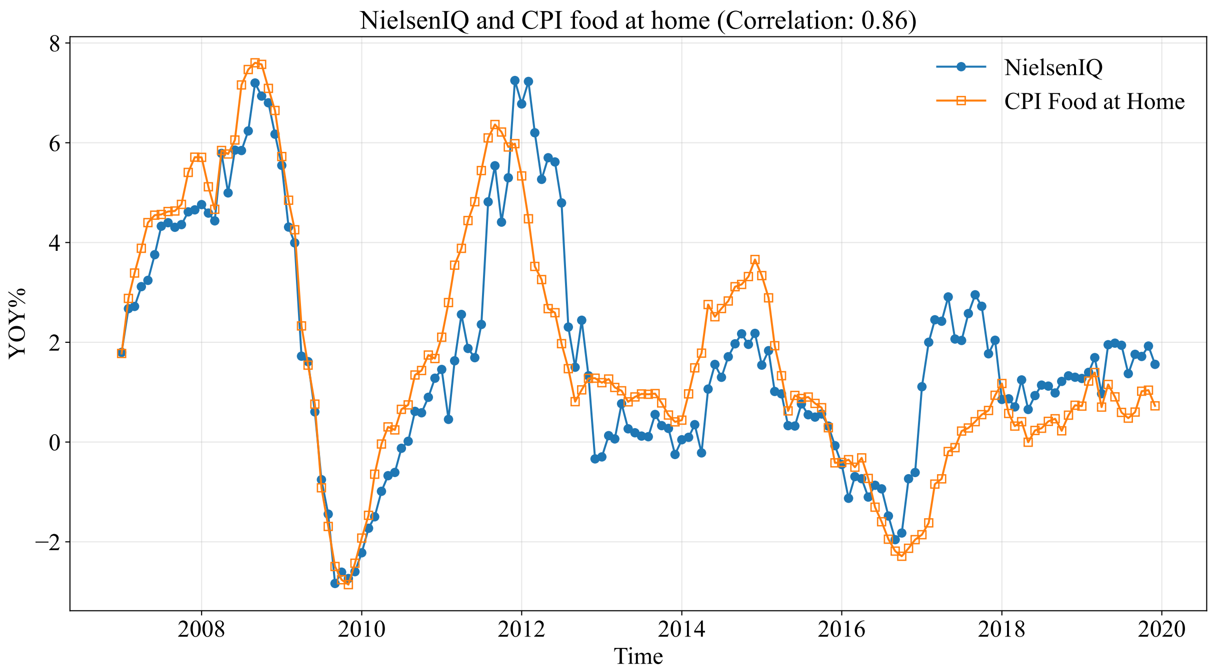
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. San Francisco Fed district.

Figure J.22: Impulse response functions to a markup shock in the San Francisco Fed district



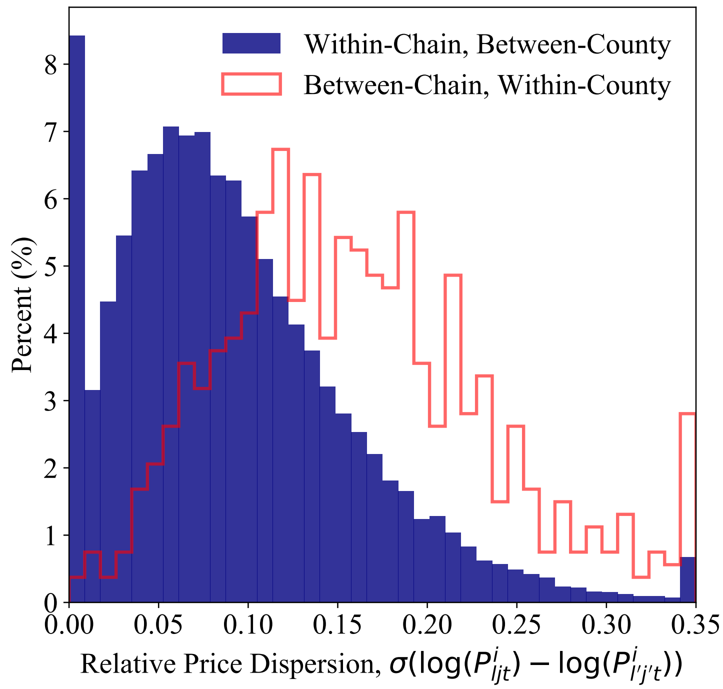
Notes: This figure plots impulse response functions for aggregate inflation and aggregate GDP to a standard deviation positive monetary shock. CIR Ratio denotes the ratio of CIR of the benchmark economy to the counterfactual economy. The counterfactual economy is an economy without multi-region firms. The counterfactual economy is an economy where firms set price in a decentralized way. The y-axis is in standard deviation of the shock. San Francisco Fed district.

Figure J.23: Price indices: NielsenIQ × CPI food at home



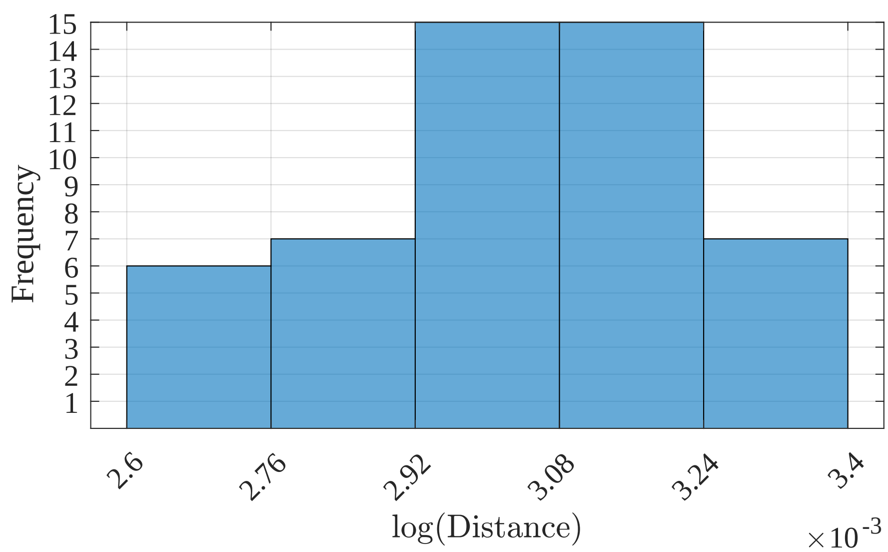
Notes: This figure plots year-over-year inflation for the NielsenIQ price index and the CPI food at home. The NielsenIQ price index was constructed following Beraja, Hurst, and Ospina (2019). It was constructed at the weekly level. To construct the monthly index, I took the average of the index within the month.

Figure J.24: Histogram of relative price dispersion



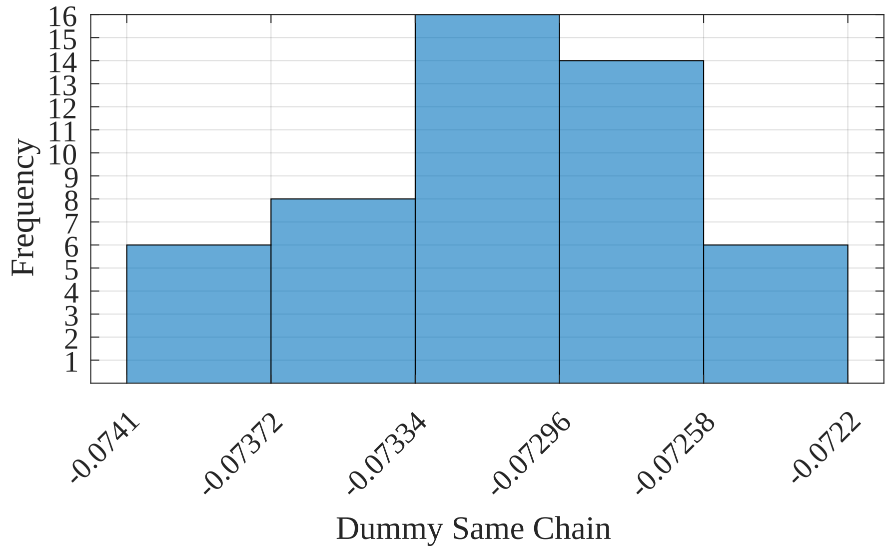
Notes: This figure plots the histogram of within-chain, between-county relative price dispersion for pairs of prices with counties ≥ 150 miles of distance and between-chain, within-county relative price dispersion.

Figure J.25: Histogram of distance coefficient



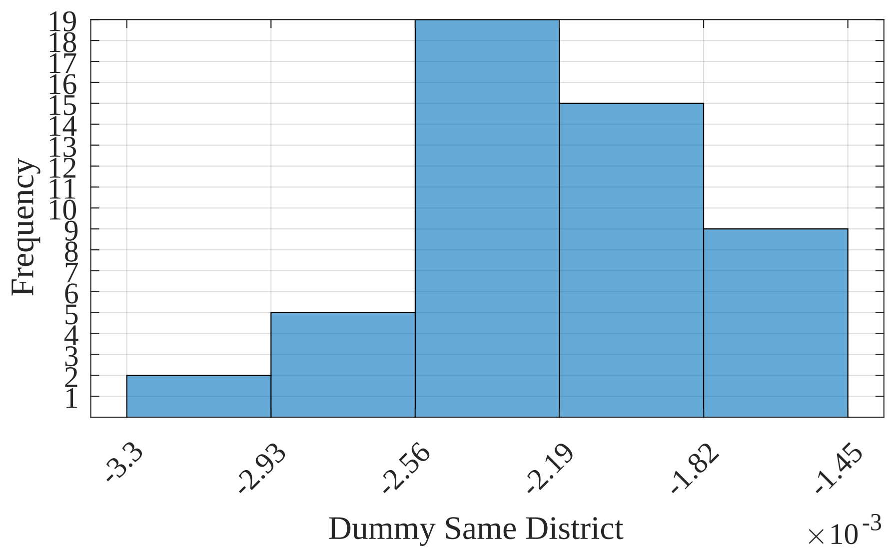
Notes:

Figure J.26: Histogram of same chain coefficient



Notes:

Figure J.27: Histogram of same district coefficient



Notes: