APPENDIX
(FOR ONLINE PUBLICATION)

A Additional Figures and Tables

Table A.1: Number of Competitors and Degree of Strategic Complementarity

	Observations	*Number of Competitors*		*Strategic Complementarity*	
		Mean	*Std. Dev.*	*Mean*	*Std. Dev.*
Industry	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Accommodation and Food Services	153	10.634	7.428	0.833	0.390
Agriculture, Forestry, and Fishing	1	32.000	.	1.000	.
Basic Chemical and Chemical Product Manufacturing	46	6.304	5.193	0.831	0.405
Beverage and Tobacco Product Manufacturing	32	9.844	6.624	0.755	0.537
Construction	206	7.083	5.415	0.858	0.383
Fabricated Metal Product Manufacturing	109	8.459	5.933	0.731	0.561
Financial and Insurance Services	413	9.075	6.672	0.809	0.440
Food Product Manufacturing	135	9.689	7.546	0.769	0.467
Furniture and Other Manufacturing	90	8.889	5.954	0.721	0.531
Information Media and Telecommunications	54	6.093	5.235	0.824	0.353
Machinery and Equipment Manufacturing	214	7.794	5.804	0.816	0.491
Non-Metallic Mineral Product Manufacturing	25	9.880	5.215	0.833	0.395
Petroleum and Coal Product Manufacturing	8	7.000	5.757	0.821	0.374
Polymer Product and Rubber Product Manufacturing	56	6.732	5.011	0.836	0.397
Primary Metal and Metal Product Manufacturing	18	6.278	5.839	0.781	0.413
Printing	58	8.621	7.684	0.824	0.421
Professional, Scientific, and Technical Services	407	7.990	6.064	0.839	0.424
Pulp, Paper and Converted Paper Product Manufacturing	16	4.875	3.739	0.873	0.359
Rental, Hiring, and Real Estate Services	121	9.702	6.196	0.837	0.485
Retail Trade	316	9.285	6.044	0.798	0.449
Textile, Leather, Clothing and Footwear Manufacturing	97	9.144	6.801	0.746	0.517
Transport Equipment Manufacturing	46	8.130	6.962	0.923	0.221
Transport, Postal, and Warehousing	197	7.746	5.458	0.860	0.482
Wholesale Trade	175	7.223	5.596	0.817	0.427
Wood Product Manufacturing	79	8.544	5.991	0.854	0.372
Total	3072	8.449	6.273	0.817	0.445

Notes: The table presents the raw (unweighted) summary statistics for the number of competitors and the degree of strategic complementarity in the survey data from New Zealand for different industries using firms' responses in the sixth and eighth waves of the survey (Coibion et al., 2018, 2021). Column (1) shows the number of observations within each industry. Columns (2) and (3) report the mean and the standard deviation of the number of competitors that firms report they face in their main product market. Columns (4) and (5) show the mean and the standard deviation for the degree of strategic complementarity from Equation (13).

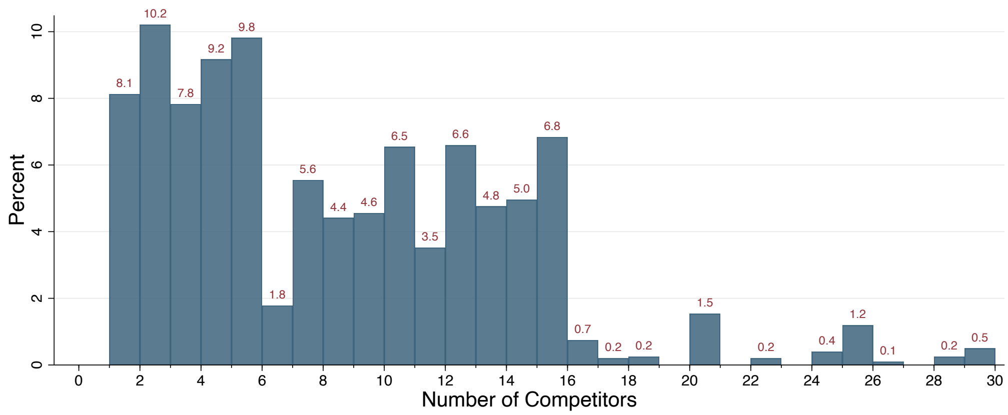


Figure A.1: Distribution of the Number of Competitors

Notes: The figure presents the raw (unweighted) distribution of the number of competitors that firms report they face in their direct product market in the sixth wave of the survey from New Zealand. The numbers over bars denote the percentage of firms within the corresponding bin. Firms with more than 30 competitors are omitted in this Figure but not in the calibration (less than 1 percent of firms report they have more than 30 competitors; the maximum number of competitors reported is 42).

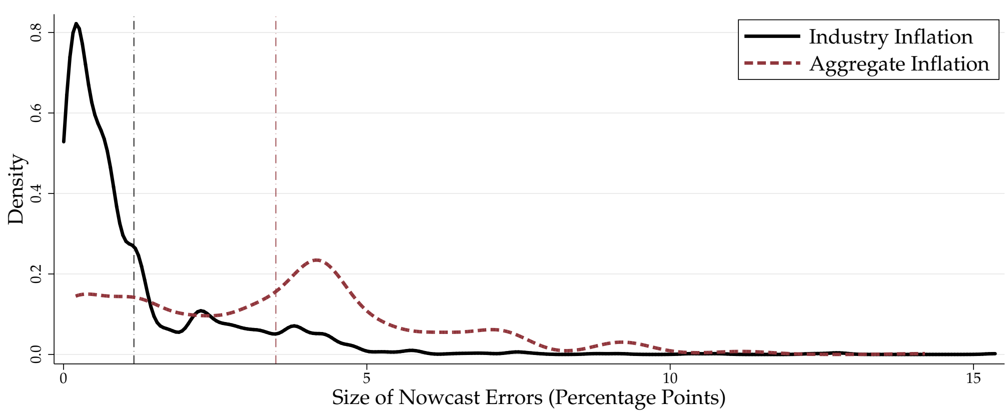


Figure A.2: Distributions of the Size of Firms' Nowcast Errors

Notes: The figure presents the raw (unweighted) distribution of the size of firms' errors in perceiving the aggregate and their industry inflation in the fourth wave of the survey (Coibion et al., 2018). The dashed vertical lines denote the means of these distributions.

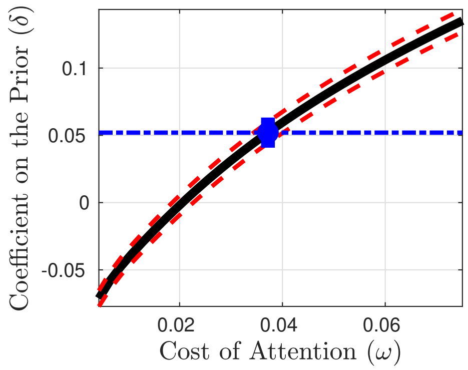


Figure A.3: Sensitivity of δ to Cost of Attention (ω)

Notes: The black line shows the predicted value of δ from the regression specified in Equation (29) in model generated data as a function of ω. The blue dot shows the equivalent estimate in the New Zealand data from Table A.2.

Table A.2: Calibration of Cost of Attention (ω)

	(1) nowcast	(2) nowcast
forecast	0.163	0.052
	(0.011)	(0.008)
perceived target		0.674
		(0.020)
Constant	3.107	0.734
	(0.102)	(0.081)
Observations	1257	1257

Notes: The table reports the result of regressing firms' nowcasts of yearly inflation on their forecasts of yearly inflation for the same horizon reported a year before. The coefficient on the lagged forecast captures the weight that firms put on their priors and increases with the degree of information rigidity. Column (1) reports the result with no controls. Column (2) controls for the firm's expectation of long-run inflation. Robust standard errors are in parentheses.

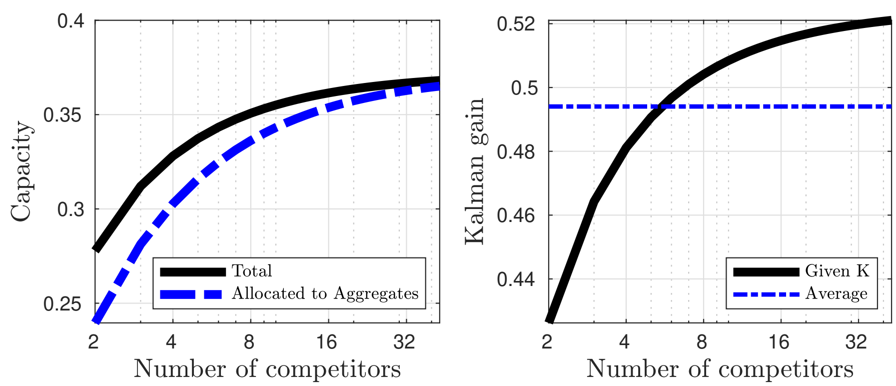


Figure A.4: Information Capacity and Kalman Gains for Different Values of K.

Notes: The left panel shows the produced information processing capacity of a firm as a function of the number of competitors within its sector in the calibrated model. The right panel shows the model implied Kalman gains of firms (weight put on the most recent signal by firms) as a function of the number of competitors within a sector. Firms with more competitors acquire more information and have larger Kalman gains. The blue dotted line shows the average Kalman gain of firms weighted by the distribution of the number of competitors in the data.

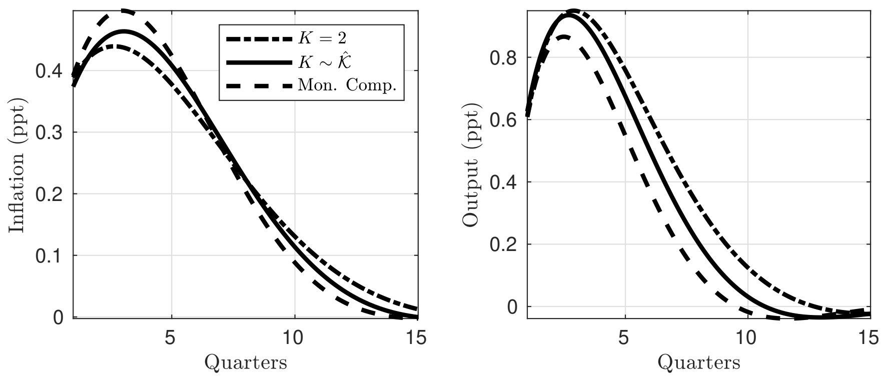


Figure A.5: IRFs to a 1% Expansionary Shock

Notes: The figure shows the impulse response functions of output and inflation to a one percent expansionary shock to the growth of nominal demand in three models. The black lines are impulse responses in the benchmark model where the distribution of the number of competitors in the model is calibrated to the empirical distribution in the data (Figure A.1). The dashed lines show the impulse responses in the model with monopolistic competition. The dash-dotted lines show the impulse responses in duopolies. Plotted impulse responses are interpolated over a finer time grid for better visual depiction.

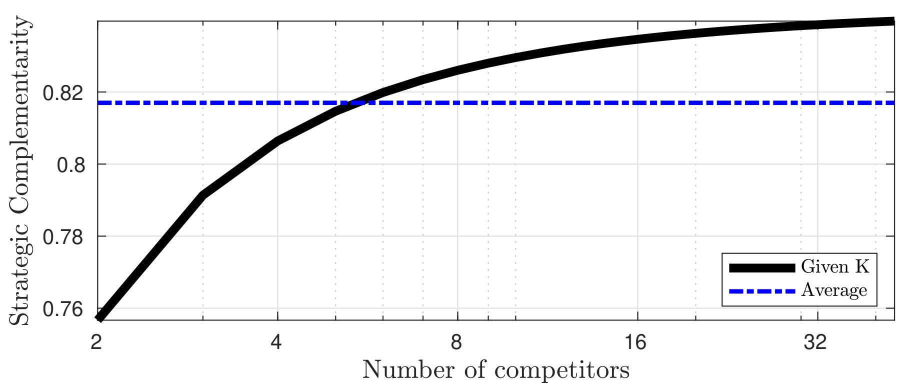


Figure A.6: Strategic complementarity as a function of K.

Notes: The figure shows the relationship between the number of competitors within a sector and the degree of strategic complementarity in pricing. Firms with a larger number of competitors have a higher degree of strategic complementarity. The dash-dotted line shows the average degree of strategic complementarity weighted by the calibrated distribution of the number of competitors to the survey data as described in Table 3.

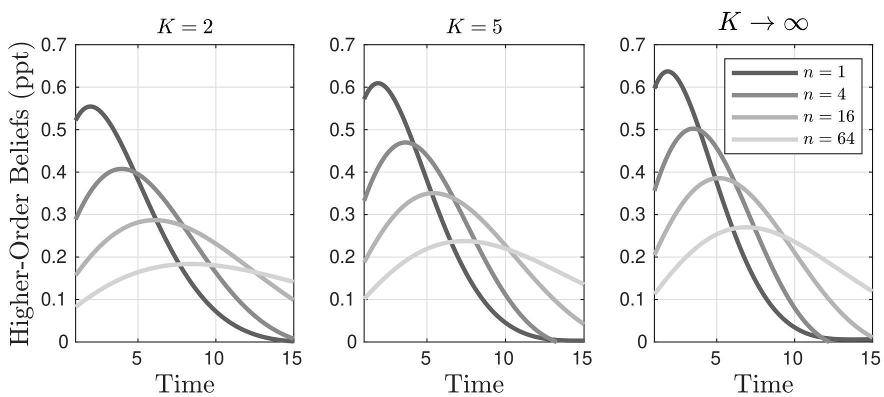


Figure A.7: IRFs of Higher-Order Beliefs to a 1% Expansionary Shock

Notes: The figure shows the IRFs of firms' higher-order beliefs to a one percent expansionary shock to the growth of nominal demand across three different models. For any given order (n), firms' n^{th} order beliefs in economies with a larger number of competitors are more responsive to the shock. This is driven by the fact that firms in more competitive economies acquire more information about the aggregate shock. Plotted impulse responses are interpolated over a finer time grid for better visual depiction.

B Mutual Information and Data Processing Inequality

In this paper, following the rational inattention literature, I use Shannon's mutual information function for measuring firms' attention. In the case of Gaussian variables, this function takes a simple form: if X and Y are two Gaussian random variables, then the mutual information between them is given by $\mathcal{I}(X;Y) = \frac{1}{2}\ln(\frac{\det(\Sigma_X)}{\det(\Sigma_{X|Y})})$, where $\Sigma_{X|Y}$ is the variance of X conditional on Y. Intuitively, the mutual information is bigger if the Y reveals more information about X, captured by a smaller $\det(\Sigma_{X|Y})$. In the other extreme, where X and Y are independent variables, $\Sigma_{X|Y} = \Sigma_X$ and the mutual information between them is zero, $\mathcal{I}(X;Y) = 0$. In other words, when X is independent of Y, then observing Y does not change the posterior of an agent about X and therefore reveals no information about X.

A result from Information Theory that I use in this Appendix is the *data processing inequality.* The following Lemma proves a weak version of this inequality for completeness.

Lemma B.1. Let $X \to Y \to Z$ be a Markov chain. Then $\mathcal{I}(X;Y) \geq \mathcal{I}(X;Z)$.

Proof. The inequality follows immediately from the chain rule of mutual information:[43]

$$\mathcal{I}(X;(Y,Z)) = \mathcal{I}(X;Y) + \mathcal{I}(X;Z|Y) = \mathcal{I}(X;Z) + \mathcal{I}(X;Y|Z)$$

Since $X \perp Z|Y$, we have $\mathcal{I}(X;Z|Y) = 0$. Thus, $\mathcal{I}(X;Y) = \mathcal{I}(X;Z) + \underbrace{\mathcal{I}(X;Y|Z)}_{\geq 0} \geq \mathcal{I}(X;Z)$. ■

C Formalizing the Static Model

This section formalizes the static game in Section 2. The Appendix is organized as follows. Appendix C.1 briefly discusses how the problem of oligopolistic firms relates to the problems studied in the previous literature and concludes by proving the feasibility and optimality of recommendation strategies for the static game with exogenous capacity. Appendix C.2 shows that for a given exogenous information processing capacity the equilibrium is unique in terms of the joint distribution of prices and the fundamental (q). Appendix C.3 derives an intuitive reinterpretation of a firm's attention problem that is discussed in Section 2. Appendix C.4 derives firms' best responses for the optimal capacity in the model with endogenous information processing capacity. Appendix C.5 derives the equilibrium values of optimal capacity in the endogenous capacity model. Appendix C.6 discusses the conditions underlying multiple equilibria with endogenous capacity and provides intuition for why multiple equilibria arise in this setting. Appendix C.7 discusses how firms' attention to the fundamental varies with the number of competitors in the model with endogenous capacity. Appendix C.8 derives first and second-order approximations to equilibrium objects in the endogenous capacity model. Finally, Appendix C.9 contains the proofs of propositions and corollaries for the static model.

[43] For a formal definition of the chain rule see Cover and Thomas (2012).

C.1. Optimal Signals in Gaussian Settings: From Single-Agent Problems to Games

This section formalizes the static game with exogenous capacity, as introduced in Section 2, and briefly discusses how well-known results for static single-agent rational inattention problems in Linear-Quadratic-Gaussian settings can be extended to this game. We start by defining the set of available signals, $\mathbb{S}$, and the set of strategies of firms in the static game, $\mathcal{A}$. The section concludes with a proposition that proves the feasibility and optimality of recommendation strategies in this game and shows that the optimal signals in the game take the well-known form of "ideal price plus noise," as in single-agent problems.

To set the stage for the analysis of the static game, consider a firm j,k's problem in Equation (1) for a given strategy of its competitors, $\varsigma \equiv (S_{l,m} \subseteq \mathbb{S}, p_{l,m}: S_{l,m} \to \mathbb{R})_{(l,m)\neq(j,k)}$:

$$\min_{S_{j,k}\subseteq\mathbb{S}} \mathbb{E}\Big[\min_{p_{j,k}:S_{j,k}\to\mathbb{R}} \mathbb{E}\Big[\Big(p_{j,k}(S_{j,k}) - p^*_{j,k}(\varsigma)\Big)^2 | S_{j,k}\Big]\Big] \quad \text{(C.1)}$$

$$s.t. \quad \mathcal{I}(S_{j,k}; q, (p_{l,m}(S_{l,m}))_{(l,m)\neq(j,k)}) \leq \kappa$$

where $p^*_{j,k}(\varsigma) = (1-\alpha)q + \alpha\sum_{l\neq k} p_{j,l}(S_{j,l}(\varsigma))$ is the ideal price of firm (j,k) given ς. Note that with the distribution of $p^*_{j,k}(\varsigma)$ given under ς, the problem above is identical to a single-agent static rational inattention problem that have been extensively studied in the literature (see, Maćkowiak, Matějka, & Wiederholt, 2023, for a review). It is known that, in such problems, optimal prices are sufficient statistics for optimal signals, and if $p^*_{j,k}(\varsigma) \sim \mathcal{N}(0,\sigma^2)$, these optimal signals are proportional to the form "ideal price plus noise:"

$$p^*_{j,k}(S^*_{j,k}) = S^*_{j,k} = \lambda p^*_{j,k}(\varsigma) + z^*_{j,k}, \quad z^*_{j,k} \perp p^*_{j,k}(\varsigma), \quad \mathbb{V}\text{ar}(z^*_{j,k}) = \lambda(1-\lambda)\sigma^2, \quad \lambda \equiv 1 - e^{-2\kappa}$$

(For such a single-agent problem see, e.g., Maćkowiak et al., 2023, Section 2.3.2).

However, extending this result to a game-theoretic setting requires an assessment of the assumptions under which it is derived: In particular, while the distribution of $p^*_{j,k}(\varsigma)$ is exogenous to the problem of the agent, it is endogenous to the game and our assumptions above on its distribution need to be derived as results in this setting. To do so, I first formally define the set $\mathbb{S}$ and the set of strategies for the game, and then state the equivalent of the result above for the game in Proposition C.1.

Definition (Rich Set of Signals)**.** Let $\mathcal{B} \equiv \{q, e_1, e_2, \ldots\}$, where q is the fundamental and e_i's are i.i.d. standard normals that are orthogonal to each other and q. Define the set $\mathbb{S}$ as the vector space of generated by $\mathcal{B}$ over the field of real numbers, i.e.,

$$\mathbb{S} \equiv \{a_0 q + \sum_{i=1}^{N} a_i e_{\sigma(i)}, N \in \mathbb{N}, (a_i)_{i=0}^{N} \subset \mathbb{R}^{N+1}, (\sigma(i))_{i=1}^{N} \subset \mathbb{N}\}. \quad \text{(C.2)}$$

Definition (Strategy Sets)**.** A strategy for firm j,k is to choose a finite vector of signals $S_{j,k} \in \mathbb{S}^{n_{j,k}}$, where $n_{j,k} \in \mathbb{N}$ is the number of signals that the firm chooses to observe, and a pricing strategy $p_{j,k}: S_{j,k} \to \mathbb{R}$

that maps the firm's signal vector to a price.[44] Thus, the set of firm j,k's pure strategies is

$$\mathcal{A}_{j,k} = \{\varsigma_{j,k} | \varsigma_{j,k} = (S_{j,k} \in \mathbb{S}^{n_{j,k}}, p_{j,k} : S_{j,k} \to \mathbb{R}), n_{j,k} \in \mathbb{N}\}.$$

Moreover, the set of pure strategies for the game is $\mathcal{A} = \{\varsigma | \varsigma = (\varsigma_{j,k})_{j,k \in J \times K}, \varsigma_{j,k} \in \mathcal{A}_{j,k}, \forall j,k \in J \times K\}$.

With the definition of $\mathbb{S}$ and strategies in $\mathcal{A}$ at hand, we can prove the following Proposition.

Proposition C.1. Suppose $\varsigma = (S_{j,k}, p_{j,k}(S_{j,k}))_{(j,k) \in J \times K} \in \mathcal{A}$ is an equilibrium. Then,

1. *Feasibility of Recommendation Strategies:* The strategy $\hat{\varsigma}(\varsigma) = (p_{j,k}(S_{j,k}), 1)_{(j,k) \in J \times K}$—where 1 denote the identity map—is also in $\mathcal{A}$.
2. *Optimality of Recommendation Strategies:* Given ς, each firm j,k is indifferent between $\hat{\varsigma}_{j,k}(\varsigma) = (p_{j,k}(S_{j,k}), 1)$ and $\varsigma_{j,k} = (S_{j,k}, p_{j,k}(S_{j,k}))$. Moreover, optimal signals under $\hat{\varsigma}_{j,k}(\varsigma)$ are proportional to the form "ideal price plus noise:"

$$p_{j,k}(S_{j,k}) = \lambda p^*_{j,k}(\varsigma) + z_{j,k}, \quad e_{j,k} \perp (q, p_{j,k}(S_{l,m}))_{(l,m) \neq (j,k)}, \mathbb{V}\text{ar}(z_{j,k}) = \lambda(\lambda - 1)\mathbb{V}\text{ar}(p)$$

Proof of Part 1: Feasibility of Recommendation Strategies. The following lemma formalizes Part 1, establishing the feasibility of recommendation strategies.

Lemma C.1. If $\varsigma = (S_{j,k} \in \mathbb{S}^{n_{j,k}}, p_{j,k} : S_{j,k} \to \mathbb{R})_{(j,k) \in J \times K} \in \mathcal{A}$ is an equilibrium, then $\forall (j,k)$, $p_{j,k}(S_{j,k})$ is itself a signal in $\mathbb{S}$.

Proof. A necessary condition for ς to be an equilibrium is that $\forall (j,k) \in (J \times L)$

$$p_{j,k}(S_{j,k}) = \text{argmin}_{p_{j,k}} \mathbb{E}[(p_{j,k} - (1-\alpha)q - \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l}(S_{j,l}))^2 | S_{j,k}].$$

which leads to the following first order condition: $p^*_{j,k}(S_{j,k}) = (1-\tilde{\alpha})\mathbb{E}[q|S_{j,k}] + \tilde{\alpha}\mathbb{E}[p^*_j(S_j)|S_{j,k}]$, where $\tilde{\alpha} \equiv \frac{\alpha + \frac{\alpha}{K-1}}{1 + \frac{\alpha}{K-1}} < 1$, and $p^*_j(S_j) \equiv K^{-1} \sum_{k \in K} p^*_{j,k}(S^*_{j,k})$. Iterating this forward, we arrive at

$$p^*_{j,k}(S_{j,k}) = \lim_{M \to \infty} ((1-\tilde{\alpha}) \sum_{m=0}^{M} \tilde{\alpha}^m \mathbb{E}^{(m)}_{j,k}[q] + \tilde{\alpha}^{M+1} \mathbb{E}^{(M+1)}_{j,k}[p^*_j(S_j)])$$

where $\mathbb{E}^{(0)}_{j,k}[q] \equiv \mathbb{E}[q|S_{j,k}]$ is firm j,k's expectation of the fundamental q, and $\forall m \geq 1$,

$$\mathbb{E}^{(m)}_{j,k}[q] = K^{-1} \sum_{l \in K} \mathbb{E}[\mathbb{E}^{(m-1)}_{j,l}[q] | S_{j,k}]$$

is firm j,k's m^{th} order higher order belief of its industry's average expectation of the fundamental. Similarly $\mathbb{E}^{(M+1)}_{j,k}[p^*_j(S_j)]$ is firm j,k's $M+1^{\text{th}}$ order belief of their industry price. Since $\tilde{\alpha} < 1$, the later term in the limit converges to zero (as long as firms' expectations of their own industry prices are not explosive under the strategy ς which is formally ruled out by Footnote 44) and we have:

$$p^*_{j,k}(S_{j,k}) = (1-\tilde{\alpha}) \sum_{m=0}^{\infty} \tilde{\alpha}^m \mathbb{E}^{(m)}_{j,k}[q]. \tag{C.3}$$

[44] As a technical assumption, we further assume that pricing strategies are L^2-integrable with respect to the measure generated by the corresponding signals—i.e., $\int |p_{j,k}(x)|^2 G(dx) < \infty$ where G is the Gaussian distribution generated by $S_{j,k} \in \mathbb{S}$. This guarantees that both unconditional and conditional expectations of firms of their competitors' prices under a given strategy is well-defined and finite.

Now, it only remains to show that $\mathbb{E}^{(m)}_{j,k}[q]$ is linear in $S_{j,k}$ for all m, which can be shown by induction. Notationwise let $\forall j,k$, let $\Sigma_{q,S_{j,k}} \equiv \mathbb{C}\text{ov}(S_{j,k},q) = \mathbb{E}[qS'_{j,k}]$. Also given j,k, $\forall l \neq k$, $\Sigma_{S_{j,l},S_{j,k}} = \mathbb{C}\text{ov}(S_{j,k},S_{j,l}) = \mathbb{E}[S_{j,l}S'_{j,k}]$ and $\Sigma_{S_{j,k}} = \mathbb{V}\text{ar}(S_{j,k}) = \mathbb{E}[S_{j,k}S'_{j,k}]$. Now, for $m=0$, $\mathbb{E}^{(0)}_{j,k}[q] = \mathbb{E}[q|S_{j,k}] = \Sigma_{q,S_{j,k}}\Sigma^{-1}_{S_{j,k}}S_{j,k}$,, which implies that 0^{th} order expectations of firms are linear in their signals. Now suppose $\forall j,l$ $\mathbb{E}^{(m)}_{j,l}[q] = A_{j,l}(m)'S_{j,l}$ for some $A_{j,l}(m) \in \mathbb{R}^{n_{j,l}}$. Thus,

$$\mathbb{E}^{(m+1)}_{j,k}[q] = K^{-1}(A_{j,l}(m) + \textstyle\sum_{l\neq k} A_{j,l}(m)\Sigma_{S_{j,l},S_{j,k}}\Sigma^{-1}_{S_{j,k}})'S_{j,k},$$
$$\implies A^{(m+1)}_{j,k} = K^{-1}(A_{j,l}(m) + \textstyle\sum_{l\neq k} A_{j,l}(m)\Sigma_{S_{j,l},S_{j,k}}\Sigma^{-1}_{S_{j,k}}) \in \mathbb{R}^{n_{j,k}}$$

which shows that the $(m+1)^{\text{th}}$ order expectation is linear in $S_{j,k}$.[45] Since this holds for all m, it follows that $p_{j,k}(S_{j,k})$ is a linear function of $S_{j,k}$. Since $\mathbb{S}$ is closed under linear combinations, this implies that $p_{j,k}(S_{j,k})$ is in $\mathbb{S}$, where

$$p_{j,k}(S_{j,k}) = \mathbb{E}[(1-\alpha)q + \alpha\tfrac{1}{K-1}\textstyle\sum_{l\neq k} p_{j,l}(S_{j,l})|S_{j,k}] \tag{C.4}$$

Furthermore, letting $p^*_{j,k}(\varsigma) = (1-\alpha)q + \alpha\frac{1}{K-1}\sum_{l\neq k}p_{j,l}(S_{j,l})$, we can see that $p^*_{j,k}(\varsigma)$ is also a linear combination of signals in $\mathbb{S}$, and thus $p^*_{j,k}(\varsigma) \in \mathbb{S}$ has a Gaussian distribution. ■

Proof of Part 2: Optimality of Recommendation Strategies. In order to confirm that recommendation strategies are optimal, we need to consider firms' deviations from an equilibrium strategy to other feasible strategies in their strategy sets. To do so, we first introduce the notation that formalizes such deviations in $\mathcal{A}_{j,k}$.

Suppose $\varsigma \in \mathcal{A}$ is an equilibrium. Let $p(\varsigma_{j,k})$ denote the optimal price of firm j,k under the given strategy (which is in $\mathbb{S}$ by Part 1). Also, let $\varsigma_{-(j,k)} \equiv \varsigma \backslash \varsigma_{j,k}$ denote the vector of the strategies for the competitors of firms j,k. Finally, let $\theta_{j,k}(\varsigma_{-(j,k)}) \equiv (q,(p(\varsigma_{j,l}))_{l\neq k},(p(\varsigma_{m,n}))_{m\neq j,n\in K})$ denote the vector of prices other than j,k augmented with the fundamental q, and define:

$$\mathbf{w} \equiv (1-\alpha, \underbrace{\tfrac{\alpha}{K-1},\ldots,\tfrac{\alpha}{K-1}}_{K-1 \text{ times}}, \underbrace{0,0,\ldots,0}_{(J-1)\times K \text{ times}}).$$

Given this notation, firm j,k's problem is to consider deviations from ς in $\mathcal{A}_{j,k}$ to solve

$$\min_{\hat{\varsigma}_{j,k}\in\mathcal{A}_{j,k}} L_{j,k}(\hat{\varsigma}_{j,k},\varsigma_{-(j,k)}) \equiv \mathbb{E}[(p(\hat{\varsigma}_{j,k}) - \mathbf{w}'\theta_{j,k}(\varsigma_{-(j,k)}))^2|S(\hat{\varsigma}_{j,k})] \tag{C.5}$$
$$s.t. \quad \mathcal{I}(S(\hat{\varsigma}_{j,k});\theta_{j,k}(\varsigma_{-(j,k)})) \leq \kappa$$

where $S(\hat{\varsigma}_{j,k})$ denotes the signals in $\mathbb{S}$ that j,k observes under the strategy $\hat{\varsigma}_{j,k}$ and given the joint distribution of $(S(\hat{\varsigma}_{j,k}),\theta_{j,k}(\varsigma_{-(j,k)}))$, the mutual information is defined in Section B. Notice that since $\mathbf{w}'\theta_{j,k}(\varsigma_{-(j,k)})$ is a linear combination of prices of other firms and the fundamental—all of which are in $\mathbb{S}$ by Part 1 of the proof—it has a Gaussian distribution that firm j,k takes as given. Thus, the problem above is identical to a single-agent rational inattention problem as in the previous literature cited at the beginning of this section, with one difference: In single agent problems, the objective is minimized over

[45] Here, I have assumed $\Sigma_{S_{j,k}}$ is invertible, which is without loss of generality: if $\Sigma_{j,k}$ is not invertible, since all signals in $S_{j,k}$ are non-zero then it must be the case that $S_{j,k}$ contains co-linear signals. In that case we can exclude the redundant signals without changing the posterior of the firm.

the set of joint distributions, whereas here, we are considering deviations in the strategy space of firm j,k. However, since we have defined this strategy space to be rich enough, the following lemma proves that these deviations are equivalent to choosing a joint distribution as in single agent problems by showing that the equilibrium strategies are weakly dominated by recomendation strategies.

Lemma C.2. For any $j,k \in J \times K$, $\forall \varsigma = (\varsigma_{j,k}, \varsigma_{-(j,k)}) \in \mathcal{A}$ that is an equilibrium, firm j,k is indifferent between $\varsigma_{j,k}$ and $\hat{\varsigma}_{j,k} = (p_{j,k}(\varsigma), 1) \in \mathcal{A}_{j,k}$. Moreover, optimal prices are proportional to the form "ideal price plus noise:"

$$p_{j,k}(\varsigma) = \lambda((1-\alpha)q + \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l}(\varsigma)) + z_{j,k}, \quad z_{j,k} \perp (q, p_{l,m}(\varsigma))_{((l,m) \neq (j,k))}$$

$$\mathbb{V}\text{ar}(z_{j,k}) = \lambda(1-\lambda)\mathbb{V}\text{ar}((1-\alpha)q + \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l}(\varsigma)), \quad \lambda = 1 - e^{-2\kappa}$$

Proof. Given $\varsigma \in \mathcal{A}$, let $\Sigma_{\varsigma_{j,k}} \equiv \mathbb{V}\text{ar}(S(\varsigma_{j,k}))$, $\Sigma_{\theta_{j,k},\varsigma_{j,k}} \equiv \mathbb{C}\text{ov}(\theta_{j,k}(\varsigma_{-(j,k)}), S(\varsigma_{j,k}))$ and $\Sigma_{\theta_{j,k}} \equiv \mathbb{V}\text{ar}(\theta_{j,k}(\varsigma_{-(j,k)}))$. Moreover, since ς is an equilibrium, then by Part 1, pricing strategies are linear and are given by:

$$p_{j,k}(\varsigma) = \mathbf{w}'\mathbb{E}[\theta_{j,k}(\varsigma_{-(j,k)})|S(\varsigma_{j,k})] = \mathbf{w}'\Sigma_{\theta_{j,k},\varsigma_{j,k}}\Sigma^{-1}_{\varsigma_{j,k}}S(\varsigma_{j,k})$$

Now, define $\hat{\varsigma}_{j,k} \equiv (p_{j,k}(\varsigma), 1)$ which is a strategy in $\mathcal{A}_{j,k}$ by Part 1 (as $p_{j,k}(\varsigma) \in \mathbb{S}$ because it is a finite linear combination of the elements of $S_{j,k}$, and $\mathbb{S}$ is rich). We have

$$L_{j,k}(\varsigma_{j,k}, \varsigma_{-(j,k)}) = \mathbf{w}'\mathbb{V}\text{ar}(\theta_{j,k}(\varsigma_{-(j,k)})|S(\varsigma_{j,k}))\mathbf{w} = \mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w} - \mathbf{w}'\Sigma_{\theta_{j,k},\varsigma_{j,k}}\Sigma^{-1}_{\varsigma_{j,k}}\Sigma'_{\theta_{j,k},\varsigma_{j,k}}\mathbf{w}.$$

so that j,k's losses under both strategies are equal:

$$L_{j,k}(\hat{\varsigma}_{j,k}, \varsigma_{-(j,k)}) = \mathbf{w}'\mathbb{V}\text{ar}(\theta_{j,k}(\varsigma_{-(j,k)})|\hat{s}_{j,k})\mathbf{w} = L_{j,k}(\varsigma_{j,k}, \varsigma_{-(j,k)}).$$

Moreover, since $p_{j,k}(\varsigma)$ is a linear function of $S_{j,k}$, $\theta_{j,k}(\varsigma_{-(j,k)}) \perp p_{j,k}(\varsigma)|S(\varsigma_{j,k})$. Therefore, by the data processing inequality in Lemma B.1, $\mathcal{I}(p_{j,k}(\varsigma); \theta_{j,k}(\varsigma_{-(j,k)})) \leq \mathcal{I}(S(\varsigma_{j,k}); \theta_{j,k}(\varsigma_{-(j,k)})) \leq \kappa$. So $\hat{\varsigma}_{j,k}$ implies the same losses as $\varsigma_{j,k}$ for firm j,k and consumes weakly less capacity. So it weakly dominates $\varsigma_{j,k}$ for firm j,k. On the other hand, $(\varsigma_{j,k}, \varsigma_{-(j,k)})$ is an equilibrium, which means that $\varsigma_{j,k}$ should weakly dominate all other strategies in $\mathcal{A}_{j,k}$, including $\hat{\varsigma}_{j,k}$. So the firm must be indifferent between the two. Also, note that the joint distribution of prices and the fundamental q is the same under both strategies.

Now, to characterize the shape of optimal signals, consider a strategy $(s_{j,k} \in \mathbb{S}, 1) \in \mathcal{A}_{j,k}$, and let $\begin{bmatrix} x^2 & \mathbf{y}' \\ \mathbf{y} & \Sigma_{\theta_{j,k}} \end{bmatrix} \equiv \mathbb{V}\text{ar}\big((s_{j,k}, \theta_{j,k}(\varsigma_{-(j,k)}))\big)$. First, recall that for $(s_{j,k} \in \mathbb{S}, 1)$ to be optimal, it has to be the case that $p_{j,k} = \mathbf{w}'\mathbb{E}[\theta_{j,k}(\varsigma_{-(j,k)})|s_{j,k}] = x^{-2}\mathbf{w}'\mathbf{y}s_{j,k}$. Thus,

$$x^2 = \mathbf{w}'\mathbf{y}.$$

Now, given $s_{j,k} \in \mathbb{S}$, the firm's loss in profits is $\mathbb{V}\text{ar}(\mathbf{w}'\theta_{j,k}(\varsigma_{-(j,k)})|s_{j,k}) = \mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w} - x^{-2}(\mathbf{w}'\mathbf{y})^2$ and the capacity constraint is $\frac{1}{2}\ln(|\mathbf{I} - x^{-2}\Sigma^{-1}_{\theta_{j,k}}\mathbf{y}\mathbf{y}'|) \geq -\kappa \Leftrightarrow x^{-2}\mathbf{y}'\Sigma^{-1}_{\theta_{j,k}}\mathbf{y} \leq \lambda \equiv 1 - e^{-2\kappa}$. Moreover, by richness of $\mathbb{S}$, we know that for any $(x, \mathbf{y})$ such that $\begin{bmatrix} x^2 & \mathbf{y}' \\ \mathbf{y} & \Sigma_{\theta_{j,k}} \end{bmatrix} \succeq 0$, there is a signal in $\mathbb{S}$ that creates this joint distribution.[46] Therefore, we let the agent choose $(x,\mathbf{y})$ freely to solve

$$\min_{(x,\mathbf{y})} \quad \mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w}-x^{-2}(\mathbf{w}'\mathbf{y})^2 \quad \text{s.t.} \quad x^{-2}\mathbf{y}'\Sigma_{\theta_{j,k}}^{-1}\mathbf{y}\leq\lambda, \quad x^2=\mathbf{w}'\mathbf{y}.$$

The solution can be derived by taking first-order conditions, but a more direct approach is to use Cauchy-Schwarz inequality $x^{-2}(\mathbf{w}'\mathbf{y})^2=x^{-2}(\Sigma_{\theta_{j,k}}^{\frac{1}{2}}\mathbf{w})'(\Sigma_{\theta_{j,k}}^{-\frac{1}{2}}\mathbf{y})\leq x^{-2}(\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w})(\mathbf{y}'\Sigma_{\theta_{j,k}}^{-1}\mathbf{y})$. Thus,

$$\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w}-x^2(\mathbf{w}'\mathbf{y})^2\geq(\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w})(1-x^{-2}\mathbf{y}'\Sigma_{\theta_{j,k}}\mathbf{y})\geq(1-\lambda)\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w},$$

We now show that our proposed signals in Part 2 of the proposition attain this global minimum. From the properties of the Cauchy-Schwarz inequality, we know it holds with equality if and only if $x^{-1}\Sigma_{\theta_{j,k}}^{-\frac{1}{2}}\mathbf{y}=c_0\Sigma_{\theta_{j,k}}^{\frac{1}{2}}\mathbf{w}$ for some constant c_0. Therefore, there is a unique vector $x^{-1}\mathbf{y}$ that attains the global minimum of the agent's problem given their constraint:$x^{-1}\mathbf{y}=c_0\Sigma_{\theta_{j,k}}\mathbf{w}$. Now, noting that the the capacity constraint should bind at the optimum (otherwise we can decrease the losses further by making signals more precise), observe that $c_0=\sqrt{\frac{\lambda}{\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w}}}$. Together with $x^2=\mathbf{w}'\mathbf{y}$, this gives us the unique $(x,\mathbf{y})$:$\mathbf{y}=\lambda\Sigma_{\theta_{j,k}}\mathbf{w},x=\sqrt{\lambda\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w}}$. Finally, notice that the set $p_{j,k}(\varsigma)_{(j,k)\in J\times K}$, as defined in Part 2 of Proposition C.1, generates these distributions—as $\mathbb{C}\text{ov}(p_{j,k}(\varsigma),\theta_{j,k}(\varsigma_{-(j,k)}))=\lambda\Sigma_{\theta_{j,k}}\mathbf{w}$, and $\mathbb{V}\text{ar}(p_{j,k}(\varsigma))=\lambda\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w}$—and all of its elements are in $\mathbb{S}$ because it is closed under finite linear operations. ■

C.2. Uniqueness of Equilibria in the Joint Distribution of Prices and the Fundamental

In Proposition C.1, we demonstrate that if a strategy is an equilibrium, then every firm is indifferent between ς and a recommendation strategy that directly proposes the implied price under ς to the firm. However, this does not guarantee the existence or uniqueness of equilibria. In this section, we address the existence and uniqueness of equilibria through the following steps. First, we show that any equilibrium $\varsigma\in\mathcal{A}$ is equivalent to an equilibrium among recommendation strategies in a sence that we precisely define below. Next, we characterize a unique equilibrium among recommendation strategies to prove uniqueness up to this equivalence relation.

Definition 1. Let $\mathcal{E}\equiv\{\varsigma\in\mathcal{A}|\varsigma \text{ is an equilibrium}\}$ denote the set of equilibria for the game. We say $\{\varsigma_1,\varsigma_2\}\subset\mathcal{E}$ are equivalent and write $\varsigma_1\sim_{\mathcal{E}}\varsigma_2$ if they imply the same joint distribution for prices of firms and the fundamental. Formally, $\varsigma_1\sim_{\mathcal{E}}\varsigma_2$ if $(q,p_{j,k}(\varsigma_1))_{j,k\in J\times K}\sim G$ if and only if $(q,p_{j,k}(\varsigma_2))_{j,k\in J\times K}\sim G$.

Note that this is clearly an equivalence relation as it satisfies reflexivity, symmetry and transitivity by properties of equality.

Lemma C.3. Suppose $\varsigma=(S_{j,k}\in\mathbb{R}^{n_{j,k}},p_{j,k}:S_{j,k}\to\mathbb{R})\in\mathcal{A}$ is an equilibrium. Then, the recommendation strategy $\hat{\varsigma}(\varsigma)=(p_{j,k}(\varsigma),1)$ as defined in Proposition C.1 is equivalent to ς: $\hat{\varsigma}(\varsigma)\sim_{\mathcal{E}}\varsigma$.

Proof. The proof is by construction. Since ς is an equilibrium it solves all firms' problems. Start from the first firm in the economy and perform the following iteration process for all firms: from previous

[46] To see why, pick $e\in\mathcal{B}$ such that $e\perp\theta_{j,k}(\varsigma_{j,-k})$. Such e exists because there are countably many infinite elements in $\mathcal{B}$ but $\theta_{j,k}(\varsigma_{-(j,k)})$ load only on finitely many of them. Then, let $s=\mathbf{y}'\Sigma_{\theta_{j,k}}^{-1}\theta_{j,k}(\varsigma_{-(j,k)})+e\sqrt{x^2-\mathbf{y}'\Sigma_{\theta_{j,k}}\mathbf{y}}$. Note that the term inside the square root is positive by positive semi-definiteness of $(x \quad \mathbf{y}';\mathbf{y} \quad \Sigma_{\theta_{j,k}})$. It is easy to verify that $(s,\theta_{j,k}(\varsigma_{-(j,k)}))$ is distributed according to this matrix.

section, we know firm $1{,}1$ has a strategy $\hat{\varsigma}_{1,1} = (s_{1,1} \in \mathbb{S}, 1)$ that is equivalent to $\varsigma_{1,1}$ given ς. Create a new strategy $\varsigma^{1,1} = (\hat{\varsigma}_{1,1}, \varsigma_{-(1,1)})$. We know that $\varsigma^{1,1}$ implies the same joint distribution as ς for the prices of all firms in the economy because we have only changed firm $1{,}1$'s strategy, and as discussed in the proof of Lemma C.2, $\hat{\varsigma}_{1,1}$ does not alter the joint distribution of prices and q. Now notice that $\varsigma^{1,1}$ is also an equilibrium because (1) firm $1{,}1$ was indifferent between $\varsigma_{1,1}$ and $\hat{\varsigma}_{1,1}$ and (2) the problem of all other firms has not changed because $1{,}1$'s price has the same joint distribution with their signals under both strategies. Now, repeat the same process for firm $1{,}2$ given $\varsigma^{1,1}$ and so on. At any step given $\varsigma^{j,k}$ repeat the process for $j,k+1$ (or $j+1,1$ if $k = K$) until the last firm in the economy. At the last step, we have $\varsigma^{J,K} = (\hat{\varsigma}_{j,k})_{j,k \in J \times K} \in \mathcal{A}$, which is (1) an equilibrium and (2) implies the same joint distribution among prices and fundamentals as ς. ■

Having shown that any equilibrium is equivalent to one among recommendation strategies, we now show that the later is unique.

Lemma C.4. Suppose the degree of strategic complementarity is strcitly less than 1, $\alpha \in [0,1)$. Then, the quotient set $\mathcal{E}/\sim_{\mathcal{E}}$ is non-empty and a singleton, i.e., all equilibria of the game are equivalent under the relationship in Definition 1.

Proof. We show this by directly characterizing the equilibrium. From the previous lemma, we know that any equilibrium is equivalent to a recommendation strategy. Suppose that $(s^*_{j,k}, 1)_{j,k \in J \times K} \in \mathcal{A}$ is an equilibrium among such strategies, and notice that in this equilibrium, every firm sets their price equal to their signal, $p_{j,k} \equiv s^*_{j,k}$. Also, Proposition C.1 showed that in this equilibrium, signals are of the following form:

$$p_{j,k} = \lambda(1-\alpha)q + \lambda\alpha \frac{1}{K-1}\sum_{l \neq k} p_{j,l} + z_{j,k},\; z_{j,k} \perp (q, p_{m,n})_{(m,n) \neq (j,k)}$$

where $\mathbb{V}\text{ar}(z_{j,t}) = \lambda(1-\lambda)\mathbb{V}\text{ar}((1-\alpha)q + \alpha\frac{1}{K-1}\sum_{l \neq k} p_{j,l})$. Now, we want to find all the joint distributions for $(q, p_{j,k})_{j,k \in J \times K}$ that satisfy this rule. Since all signals are Gaussian, the joint distributions will also be Gaussian.

To derive this distribution, we start by characterizing the covariance of any firm's price with the fundamental. For any industry j, let $p_j \equiv (p_{j,k})_{k \in K}$ and $z_j \equiv (z_{j,k})_{k \in K} \perp q$. Moreover, for ease of notation, in this section, let $\gamma \equiv \frac{1}{K-1}$. Now, the equilibrium condition implies $p_j = \lambda(1-\alpha)\mathbf{1}q + \lambda\alpha\gamma(\mathbf{1}\mathbf{1}' - \mathbf{I})p_j + z_j$ where $\mathbf{1}$ is the unit vector in $\mathbb{R}^K$, and $\mathbf{I}$ is identity matrix in $\mathbb{R}^{K \times K}$ (therefore $\mathbf{1}\mathbf{1}' - \mathbf{I}$ is a matrix with zeros on diagonal and 1's elsewhere). With some algebra it is straightforward to show that $\mathbb{C}\text{ov}(p_j, q) = \frac{\lambda - \lambda\alpha}{1 - \lambda\alpha}\mathbf{1}$. Thus, in any equilibrium, the covariance of any firm's price with the fundamental q has be to equal to

$$\delta \equiv \frac{\lambda - \lambda\alpha}{1 - \lambda\alpha} \tag{C.6}$$

Next, we show that the prices of any two firms in separate industries are orthogonal conditional on the fundamental. Let p_j be the vector of prices in industry j as defined above. Pick any firm from any other industry $l,m \in J\times K, l\neq j$. Notice that by the equilibrium conditions z_j is orthogonal to $p_{l,m}$. Now, notice that

$$\mathbb{C}\text{ov}(p_j,p_{l,m}) = \lambda(1-\alpha)\mathbf{1}\underbrace{\mathbb{C}\text{ov}(q,p_{l,m})}_{=\delta} + \lambda\alpha\gamma(\mathbf{1}\mathbf{1}'-\mathbf{I})\mathbb{C}\text{ov}(p_j,p_{l,m}) + \underbrace{\mathbb{C}\text{ov}(z_j,p_{l,m})}_{=0}.$$

With some algebra, we get $\mathbb{C}\text{ov}(p_j,p_{l,m}) = \delta^2\mathbf{1} \Rightarrow \mathbb{C}\text{ov}(p_j,p_{l,m}|q) = 0$. Therefore, in any equilibrium, prices of any two firms in two different industries are only correlated through the fundamental. This implies that firms do not pay attention to mistakes of firms in other industries.

Now, we only need to derive the joint distribution of prices within industries. We have $p_j = \mathbf{B}(\lambda(1-\alpha)\mathbf{1}q + z_j)$ where $\mathbf{B} \equiv \frac{1}{1+\alpha\lambda\gamma}\mathbf{I} + \frac{\alpha\lambda\gamma}{(1+\alpha\lambda\gamma)(1-\alpha\lambda)}\mathbf{1}\mathbf{1}'$. This gives $p_j = \delta\mathbf{1}q + \mathbf{B}z_j$, where $\mathbf{B}z_j \perp q$. This corresponds to the decomposition of the prices of firms to parts that are correlated with the fundamental and their mistakes. The vector $\mathbf{B}z_j$ is the vector of firms' mistakes in industry j, and is the same as the vector v_j in the text. Let $\Sigma_{z,j} = \mathbb{C}\text{ov}(z_j,z_j)$ and $\Sigma_{p,j} = \mathbb{C}\text{ov}(p_j,p_j)$. We have $\Sigma_{p,j} = \delta^2\mathbf{1}\mathbf{1}' + \mathbf{B}\Sigma_{z,j}\mathbf{B}'$. Also, since $z_{j,k} \perp p_{j,l\neq k}$, we have $\mathbf{D}_j \equiv \mathbb{C}\text{ov}(p_j,z_j) = \mathbf{B}\Sigma_{z,j}$ where $\mathbf{D}_j$ is a diagonal matrix whose k'th element on the diagonal is $\mathbb{V}\text{ar}(z_{j,k})$. From the equilibrium conditions we have

$$\begin{aligned}\mathbb{V}\text{ar}(z_{j,k}) &= \lambda(1-\lambda)\mathbb{V}\text{ar}((1-\alpha)q + \alpha\gamma\sum_{l\neq k}p_{j,l}) \\ &= \lambda(1-\lambda)(1-\alpha)^2 + \lambda(1-\lambda)\alpha^2\gamma^2\mathbf{w}_k'\Sigma_{p,j}\mathbf{w}_k + 2\lambda(1-\lambda)\alpha(1-\alpha)\delta\end{aligned}$$

where $\mathbf{w}_k$ is a vector such that $\mathbf{w}_k' p_j = \sum_{l\neq k} p_{j,l}$. This gives K linearly independent equations and K unknowns in terms of the diagonal of $\mathbf{D}_j$. Guess that the unique solution to this is symmetric. After some algebra, we get that the implied distribution for prices is such that

$$\mathbb{V}\text{ar}(p_{j,k}) = \frac{1-\alpha\lambda}{1-\alpha\tilde{\lambda}}\lambda^{-1}\delta^2, \forall j,k; \mathbb{C}\text{ov}(p_{j,k},p_{j,l}) = \frac{1-\alpha\lambda}{1-\alpha\tilde{\lambda}}\frac{\tilde{\lambda}}{\lambda}\delta^2, \forall j,k,l\neq k, \tag{C.7}$$

where $\tilde{\lambda} \equiv \frac{\lambda+\alpha\gamma\lambda}{1+\alpha\gamma\lambda}$.

Thus, any equilibrium should have the same distribution of prices and fundamentals derived in this proof, which concludes our proof of existence and uniqueness. ■

C.3. Reinterpretation of a Firm's Attention Problem

This section presents an alternative formulation of the firms' attention problem, where they maximize their payoffs by choosing the correlation of their prices with the fundamental and the mistakes of their competitors. This alternative formulation is mathematically equivalent to the equilibrium characterized in the previous section, but provides additional economic insights about firms' incentives.

Take any firm $j,k \in J\times K$ and suppose all other firms in the economy are playing the equilibrium strategy. Moreover, here I take it as given that the firm does not pay attention to mistakes of firms in other industries ($\mathbb{C}\text{ov}(p_{j,k},p_{l,m}|q)_{l\neq j} = 0$). Now, take a strategy $\varsigma_{j,-k}$ for other firms and decompose the average price of others under this strategy to its projection on q and the part that is orthogonal to q: $p_{j,-k}(\varsigma_{j,-k}) = \frac{1}{K-1}\sum_{l\neq k} p_{j,l}(\varsigma_{j,l}) = \delta q + v_{j,-k}$. Furthermore, let $\sigma_v^2 \equiv \mathbb{V}\text{ar}(v_{j,-k})$ be the variance of the average mistake of other firms in j,k's industry when they play $\varsigma_{j,-k}$. For any $s_{j,k} \in \mathbb{S}$ define $\rho_q(s_{j,k}) \equiv$

$cor(s_{j,k},q), \rho_v(s_{j,k}) \equiv cor(s_{j,k},v_{j,-k})$. Notice that firm j,k's loss in profit, given that it observes $s_{j,k}$, is

$$\mathbb{V}\text{ar}((1-\alpha)q+\alpha p_{j,-k}|s_{j,k}) = (1-\alpha+\alpha\delta)^2\mathbb{V}\text{ar}(q+\frac{\alpha}{1-\alpha(1-\delta)}v_{j,-k}|s_{j,k}).$$

With some algebra, it is straightforward to show that the variance in the second part of the above equation is given by

$$\mathbb{V}\text{ar}(q+\frac{\alpha}{1-\alpha(1-\delta)}v_{j,-k}|s_{j,k}) = 1+(\frac{\alpha}{1-\alpha(1-\delta)})^2\sigma_v^2-(\rho_q(s_{j,k})+\frac{\alpha\sigma_v}{1-\alpha(1-\delta)}\rho_v(s_{j,k}))^2.$$

Now, to derive the information constraint in terms of the two correlation terms, we have

$$\mathcal{I}(s_{j,k};(q,p^*_{j,-k}))\leq\kappa \Leftrightarrow \frac{1}{2}\ln(\frac{\mathbb{V}\text{ar}(s_j)}{\mathbb{V}\text{ar}(s_{j,k}|(q,p^*_{j,-k}))})\leq\kappa$$

Notice that $\frac{\mathbb{V}\text{ar}(s_j|(q,p^*_{j,-k}))}{\mathbb{V}\text{ar}(s_j)} = 1-(\rho_q(s_j)^2+\rho_v(s_j)^2)$. Thus, the information constraint becomes$\rho_q^2(s_j)+\rho_v^2(s_j)\leq\lambda\equiv 1-e^{-2\kappa}$. So, j,k's problem reduces to

$$\max_{\rho_q,\rho_v}(\rho_q(s_{j,k})+\frac{\alpha\sigma_v}{1-\alpha(1-\delta)}\rho_v(s_{j,k}))^2 \quad s.t. \quad \rho_q(s_{j,k})^2+\rho_v(s_{j,k})^2\leq\lambda.$$

C.4. Derivations for Optimal Information Processing Capacity

In this section, we establish conditions for the optimal choice of an endogenous κ, as formalized and discussed in Section 2.4. The section concludes by deriving the best response function of firms for their optimal capacities.

Consider a strategy for firm j,k's competitors where $p_{j,-k}=\delta q+v_{j,-k}, v_{j,-k}\sim\mathcal{N}(0,\sigma_v^2), v_{j,-k}\perp q$. Then, given any $\kappa_{j,k}$, and letting $\lambda_{j,k}\equiv 1-e^{-2\kappa_{j,k}}$, we know from Equation (2) that the optimal strategy of the firm is such that

$$\begin{aligned} p_{j,k} = S_{j,k} &= \lambda_{j,k}p^*_{j,k}+z_{j,k} \\ &= \lambda_{j,k}(1-\alpha+\alpha\delta)q+\alpha\lambda_{j,k}v_{j,-k}+z_{j,k} \end{aligned}$$

where, $p^*_{j,k}\equiv(1-\alpha+\alpha\delta)q+\alpha v_{j,-k}$ is firm j,k's ideal price under its competitors' strategy. We also know form Equation (2) that if

$$V^*_{j,-k}\equiv\mathbb{V}\text{ar}(p^*_{j,k}) = (1-\alpha+\alpha\delta)^2+\alpha^2\sigma_v^2$$

is the unconditional variance of $p^*_{j,k}$, then $z_{j,t}\sim\mathcal{N}(0,\lambda_{j,k}(1-\lambda_{j,k})V^*_{j,-k})$. Thus, calculating the firm's expected losses from mis-pricing under $\kappa_{j,k}$ we have

$$\begin{aligned} \frac{1}{2}B\mathbb{E}[(p_{j,k}-p^*_{j,k})^2|S_{j,k}] &= \frac{1}{2}B(1-\lambda_{j,k})V^*_{j,-k} \\ &= \frac{1}{2}e^{-2\kappa_{j,k}}BV^*_{j,-k} \end{aligned}$$

Now, replacing this into the objective of the firm we arrive at the following problem:

$$\min_{\kappa_{j,k}\geq 0}\{\frac{1}{2}e^{-2\kappa_{j,k}}BV^*_{j,-k}+\omega\kappa_{j,k}\}$$

Now, when the constraint does not bind, the first-order condition of this problem gives us the optimal

$\kappa_{j,k}$ as:

$$-e^{-2\kappa_{j,k}}BV^*_{j,-k}+\omega=0$$
$$\Rightarrow\kappa_{j,k}=\frac{1}{2}\ln(\frac{BV^*_{j,-k}}{\omega})$$

which is strictly positive in accordance with the constraint when $BV^*_{j,-k}>\omega$. Thus,

$$\kappa_{j,k}=\frac{1}{2}\max\{0,\ln(BV^*_{j,-k}/\omega)\}=\begin{cases}\frac{1}{2}\ln(\frac{BV^*_{j,-k}}{\omega}) & BV^*_{j,-k}>\omega\\ 0 & BV^*_{j,-k}\leq\omega\end{cases}$$

C.5. Derivation of $V^*_{j,-k}$ in a Symmetric Equilibrium

In this section, we use the best responses of firms for endogenous capacity from above to derive the value of this endogenous capacity in a symmetric equilibrium as function of the underlying parameters of the model.

In the previous section, we saw that the firms' best responses for their optimal information capacities depend on the variance of their ideal prices, $V^*_{j,-k}$, which in turn depends on the strategies of their competitors. Here, we first characterize the value of $V^*_{j,-k}$ in a symmetric equilibrium, which then allows us to derive the optimal information processing capacity of firms in a symmetric equilibrium.

Consider a symmetric equilibrium where a firm j,k's competitors' strategy is such that $p_{j,-k}=\delta q+v_{j,-k}, v_{j,-k}\sim\mathcal{N}(0,\sigma_v^2), v_{j,-k}\perp q$, for some δ and σ_v. Thus, dropping the $j,-k$ index from $V^*_{j,-k}$, we have

$$\begin{aligned}V^*&=\mathbb{V}\text{ar}((1-\alpha)q+\alpha p_{j,-k})\\&=\mathbb{V}\text{ar}((1-\alpha+\alpha\delta)q+\alpha v_{j,-k})\\&=(1-\alpha+\alpha\delta)^2+\alpha^2\sigma_v^2\end{aligned}\tag{C.8}$$

Now, we need to calculate σ_v^2, the variance of the average mistakes of a firm's competitors in the symmetric equilibrium. For that, we need to derive the relationship between firms' mistakes in the equilibrium. Given firm j,k's optimal attention strategy and given the optimal $\kappa^*=\kappa^*_{j,k}$ that we solved above, we know from Equation (2) that:

$$p_{j,k}=\lambda((1-\alpha+\alpha\delta)q+v_{j,-k})+z_{j,k},\quad z_{j,k}\perp(q,v_{j,-k}),\quad z_{j,k}\sim\mathcal{N}(0,\lambda(1-\lambda)V^*)\tag{C.9}$$

where $\lambda=1-e^{-2\kappa^*}$. A similar decomposition of firm j,k's price in the symmetric equilibrium to its projection on q and a mistake $v_{j,k}$ gives

$$p_{j,k}=\hat{\delta}q+v_{j,k}=\lambda(1-\delta+\alpha\delta)q+\lambda\alpha v_{j,-k}+z_{j,k}\tag{C.10}$$

By symmetry, $\delta=\hat{\delta}$ and we have $\delta=\frac{(1-\alpha)\lambda}{1-\alpha\lambda}$ as before. Also, recalling the parameter $\gamma=\frac{1}{K-1}$, we can write:

$$v_{j,k}=\alpha\lambda v_{j,-k}+z_{j,k},\tag{C.11}$$

$$v_{j,-k}=\gamma\sum_{l\neq k}v_{j,l}\tag{C.12}$$

Thus, to calculate σ_v^2, the variance of $v_{j,-k}$, we need to know the covariance matrix of the vector of mistakes, $(v_{j,l})_{l\in[K]}$. However, note that due to symmetry, this covariance matrix is summarized by two parameters: the variance of each of $v_{j,k}$'s and the covariance of any two mistakes $v_{j,k}$ and $v_{j,l\neq k}$. To calculate these two objects, first we take the variance of the two sides in Equation (C.11):

$$\mathbb{V}\text{ar}(v_{j,k}) = \alpha^2\lambda^2\sigma_v^2 + \lambda(1-\lambda)V^* \tag{C.13}$$

Second, we take the covariance of both sides of Equation (C.11) with $v_{j,-k}$:

$$\mathbb{C}\text{ov}(v_{j,k}, v_{j,-k}) = \alpha\lambda\sigma_v^2 \tag{C.14}$$

which is implied by $z_{j,k} \perp v_{j,-k}$. Now, the final step is to see that

$$\mathbb{C}\text{ov}(v_{j,k}, v_{j,-k}) = \gamma\sum_{l\neq k}\mathbb{C}\text{ov}(v_{j,k}, v_{j,l}) = \mathbb{C}\text{ov}(v_{j,k}, v_{j,l\neq k}), \forall l\neq k \Rightarrow \mathbb{C}\text{ov}(v_{j,k}, v_{j,l\neq k}) = \alpha\lambda\sigma_v^2 \tag{C.15}$$

Taking the variance of both sides of Equation (C.12) with $v_{j,l\neq k}$, and using Equations (C.13) and (C.15) we get:

$$\begin{aligned}\sigma_v^2 &= \gamma\mathbb{V}\text{ar}(v_{j,k}) + (1-\gamma)\mathbb{C}\text{ov}(v_{j,k}, v_{j,l\neq k}) \\ &= \gamma\alpha^2\lambda^2\sigma_v^2 + \gamma\lambda(1-\lambda)V^* + (1-\gamma)\alpha\lambda\sigma_v^2 \\ &= \frac{\gamma\lambda(1-\lambda)V^*}{(1-\alpha\lambda)(1+\gamma\alpha\lambda)}\end{aligned} \tag{C.16}$$

Combining this equation with Equation (C.8) and using $\delta = \frac{(1-\alpha)\lambda}{1-\alpha\lambda}$, we get:

$$\begin{aligned}V^* &= (\frac{1-\alpha}{1-\alpha\lambda})^2\mathbb{V}\text{ar}(q) + \alpha^2\frac{\gamma\lambda(1-\lambda)}{1-\gamma\alpha^2\lambda^2-(1-\gamma)\alpha\lambda}V^* \\ &= (\frac{1-\alpha}{1-\alpha\lambda})^2\frac{1+\gamma\alpha\lambda}{1+\gamma\alpha\frac{(1-\alpha)\lambda}{1-\alpha\lambda}}\mathbb{V}\text{ar}(q) \\ &= \frac{(1-\alpha)^2(K-1+\alpha\lambda)}{(1-\alpha\lambda)^2(K-1)+\alpha(1-\alpha)(1-\alpha\lambda)\lambda}\mathbb{V}\text{ar}(q)\end{aligned} \tag{C.17}$$

where we have normalized $\mathbb{V}\text{ar}(q) = 1$.

C.6. Discussion of Multiple Symmetric Equilibria with Endogenous Capacity

In this section, I discuss the possibility of multiple symmetric equilibria in the model with endogenous capacity. The section concludes by discussing when multiple equilibria might arise under certain values of parameters and establishes conditions on the parameter space for uniqueness.

Since we have shown in Appendix C.2 that for any fixed κ, the model has a unique symmetric equilibrium, we can conclude that the model has multiple symmetric equilibria if and only if there are multiple values of κ^* (or $\lambda^* \equiv 1 - e^{-2\kappa^*}$) that satisfy the equilibrium conditions. As derived in Appendices C.4 and C.5 and shown in Equations (9) and (10), in any symmetric equilibrium, λ^* is a solution to the following two equations:

$$\lambda^* = \max\{0, 1 - \frac{\omega}{BV^*}\} \tag{C.18}$$

$$V^* = (\frac{1-\alpha}{1-\alpha\lambda^*})^2\frac{K-1+\alpha\lambda^*}{K-1+\alpha\lambda^*\frac{1-\alpha}{1-\alpha\lambda^*}}\underbrace{\mathbb{V}\text{ar}(q)}_{=1} \tag{C.19}$$

To see whether multiple λ^*'s satisfy these equations, it is useful to consider both of these equations in the (V^*,λ^*) plane and investigate their intersections. It is straightforward to see that both equations define λ^* as a weakly increasing function of V^*. Normalizing $\mathbb{V}\text{ar}(q)=1$ as in the main text, the first equation takes values of $V^* \in [0,1]$ and maps it to a range of $\lambda^* \in [0,1-\frac{\omega}{B}]$ while the second equation takes values of $\lambda^* \in [0,1]$ and maps it to a range of $V^* \in [(1-\alpha)^2,1]$.

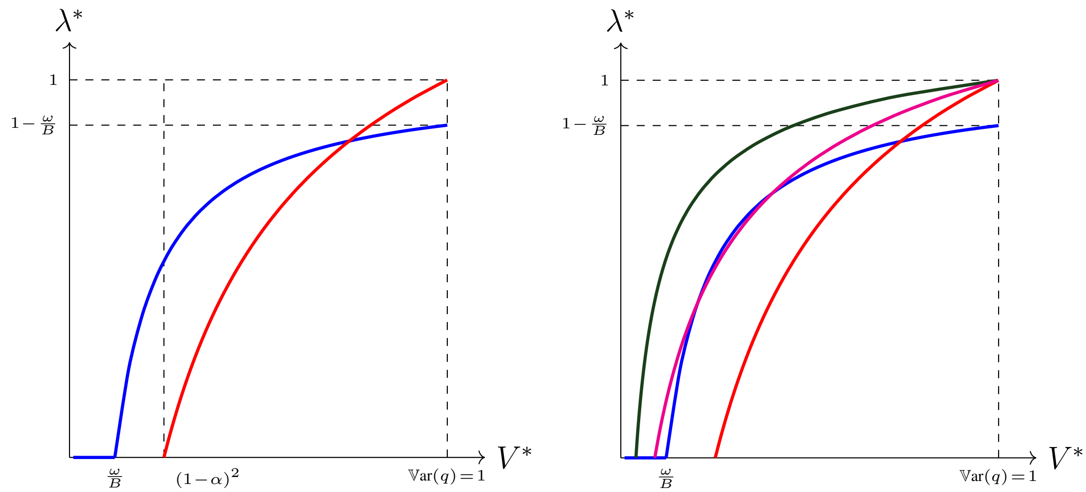


Figure C.1: Symmetric Equilibria with Endogenous Information Processing Capacity

Notes: The figure shows how the equilibrium λ^* is determined by the intersection of the two curves defined by Equations (9) and (10). In the left panel, the blue curve depicts Equation (10) with $\omega/B=0.15$, and the red curve depicts Equation (9) with $\alpha=0.5$. In the right panel, the blue curve is is kept the same, but we have added two additional curves that depict Equation (9) for two additional values of $\alpha=0.8$ (dark green) and $\alpha=0.7$ (magenta). When $(1-\alpha)^2>\omega/B$, there is a unique equilibrium with $\lambda^*>0$ (red curve), but when $(1-\alpha)^2<\omega/B$, there is always an equilibrium with $\lambda^*=0$ (both dark green and magenta curves) and possibly other equilibria with $\lambda^*>0$ (magenta curve).

As proved formally in Proposition 4, it follows that if $(1-\alpha)^2>\omega/B$, these two curves only intersect once, guaranteeing the uniqueness of a symmetric equilibrium in which $\lambda^*>0$.[47] An example of this case is depicted in the left panel of Figure C.1. On the other hand, if $(1-\alpha)^2<\omega/B$, the two curves intersect at least once at $\lambda^*=0$ (the dark green curve in the right panel of Figure C.1), and possibly at one or two other points with $\lambda^*>0$. The magenta curve in the right panel Figure C.1 shows an example where there are three equilibria, one with $\lambda^*=0$, and two with $\lambda^*>0$.

Intuition for Multiple Equilibria. To see why multiple equilibria arise in terms of information processing capacity, it is useful to revisit the economic incentives of firms in information acquisition. The key force here is that strategic complementarities in pricing induce strategic complementarities in information acquisition.

[47] See Proposition 4 and its proof for a precise argument. A brief version of this argument in this setting is as follows: the second equation defines an increasing curve that connects $(V^*,\lambda^*)=((1-\alpha)^2,0)$ and $(V^*,\lambda^*)=(1,1)$. Moreover, if $(1-\alpha)^2>\omega/B$ then first equation is strictly increasing and concave in the domain $V^*\in[(1-\alpha)^2,1]$. Thus, there can only be at most one crossing between the two curves. Moreover, since the first equation ranges from $\lambda^*=0$ to $\lambda^*=1-\omega/B$ there should at least be one crossing between the two curves. Hence, with $(1-\alpha)^2>\omega/B$ there is a unique equilibrium with $\lambda^*>0$.

For example, consider a situation where all of a firm j,k's competitors choose a zero capacity, $\lambda^*_{j,-k}=0$. In this case, their prices will not respond to any fundamental shocks and will be fixed at their prior expected value of q, which has been normalized to 0. In this situation, firm j,k has no incentive to learn about its competitors' prices because they will not respond to any shocks. The only incentive for firm j,k to acquire information is to learn about q, which only affects their ideal price with a weight of $1-\alpha$. Thus, the larger α is, the less valuable information about q becomes. If α or ω are large enough such that the benefit of acquiring information about q is not greater than the cost, then firm j,k will also choose a zero capacity, resulting in a symmetric equilibrium with $\lambda^*=0$.

However, if firm j,k's competitors follow a strategy with a positive λ^*, then even with all other parameters fixed, firm j,k will have a higher incentive to acquire information. This is because, in addition to the incentive to learn about q, the firm also values information about its competitors' prices, which now partially respond to shocks and may be subject to mistakes. This is depicted by the magenta curve in the right panel of Figure 1, where α and ω are in a range where both types of equilibria can coexist.

Finally, if ω or α are large enough, no firm will ever choose a positive capacity independent of what its competitors do, resulting in $\lambda^*=0$ as the only possible equilibrium (dark green curve in the right panel of Figure C.1). On the other hand, if ω or α are small enough, the value of information, even when competitors choose a zero capacity, is high enough relative to its cost that any firm will always choose a positive capacity regardless of what its competitors do (red curve in the right panel of Figure C.1). See Proposition 4 and its proof for a formal statement of the latter case.

C.7. Attention to Fundamental with Endogenous Capacity

This section extends the predictions of Proposition 1 to the case with endogenous capacity. Specifically, we examine how firms' attention to fundamental shocks depends on the parameters of the model when firms choose their information capacities as in Section 2.4. As we have already shown that the symmetric equilibrium is unique for a given κ in Appendix C.2, our focus is on studying how the correlation between the fundamental and firms' signals varies with model parameters under different values of κ that can arise in a symmetric equilibrium with endogenous capacities.

Using Equations (C.6) and (C.7) in Appendix C.2, for a given λ^* in a symmetric equilibrium, the correlation between a firm's price and the fundamental is given by

$$\rho_q^{*2}=\frac{K-1+\alpha\frac{(1-\alpha)\lambda^*}{1-\alpha\lambda^*}}{K-1+\alpha\lambda^*}\lambda^* \tag{C.20}$$

Thus, we can see that ρ_q^{*2} depends on K and α both directly—i.e., holding λ^* fixed as discussed in Proposition 1—and indirectly through λ^*. In general, either one of these forces can dominate the other but for small enough values of ω the predictions of Proposition 1 persist. In the remainder of this section, I provide some intuition and two examples in each of which a different force dominates. In the next section, I do a Taylor expansion of ρ_q^{*2} around a small ω and show that the predictions of Proposition 1 are indeed valid for small ω.

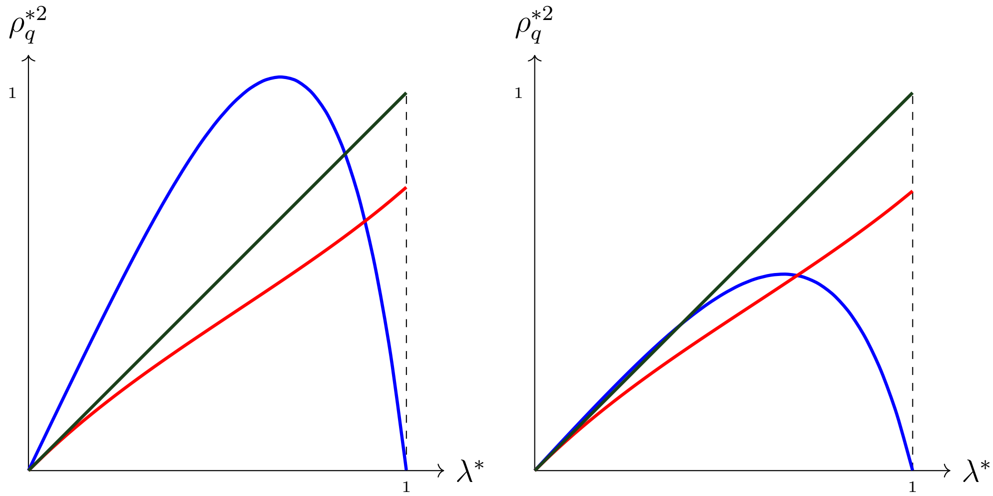


Figure C.2: Firms' Attention to the Fundamental q

Notes: The figure shows two different parameterizations of Equation (C.20) (the red curve with $K=2$ and the green curve with $K\to\infty$) and Equation (C.24) (the blue curve). The figure on the left-hand side shows the joint determination of ρ_q^{*2} when ω is small (I have set $\omega/B=.12$ and $\alpha=0.5$). In this case, the intersection for both $K=2$ and $K=\infty$ are on the decreasing side of the blue curve and thus as K increases ρ_q^{*2} increases (i.e., the force discussed in Proposition 1 persists) even though λ^* decreases (per Proposition 4). The figure on the right-hand side shows the joint determination of ρ_q^{*2} when ω is large (I have set $\omega/B=.25$ and $\alpha=0.48$). In this case, the intersection for $K\to\infty$ is on the increasing side of the blue curve and thus as K increases ρ_q^{*2} decreases (i.e., the force discussed in Proposition 1 is reversed by the decline in λ^*).

Perhaps the best way to see how these two effects interact is to consider the equation above in conjunction with the equations that implicitly characterize λ^*. In particular, recall that λ^* is given by

$$\lambda^*=\max\{0,1-\frac{\omega}{BV^*}\} \tag{C.21}$$

$$V^*=(\frac{1-\alpha}{1-\alpha\lambda^*})^2\frac{K-1+\alpha\lambda^*}{K-1+\alpha\lambda^*\frac{1-\alpha}{1-\alpha\lambda^*}} \tag{C.22}$$

Now consider an equilibrium with $\lambda^*>0$ (otherwise $\rho_q^*=0$ and does not vary with neither α nor K). Replacing V^* in the first equation with the second equation and using the expression for ρ_q^{*2} from above, we obtain

$$\lambda^*=1-\frac{\omega}{BV^*}=1-\frac{\omega}{B\lambda^*}(\frac{1-\alpha\lambda^*}{1-\alpha})^2\rho_q^{*2} \tag{C.23}$$

Re-writing this equation such that ρ_q^{*2} is on the left-hand side, we have

$$\rho_q^{*2}=\frac{B\lambda^*(1-\lambda^*)}{\omega}(\frac{1-\alpha}{1-\alpha\lambda^*})^2 \tag{C.24}$$

Therefore, the pair (λ^*,ρ_q^{*2}) in the unique equilibrium are characterized by the intersection of two curves defined by Equations (C.20) and (C.24). Figure C.2 depicts two different parameterizations of these equations, where on the left-hand side panel, where ω/B is small, an increase in K increases ρ_q^{*2} (firms' attention to fundamental increases with K) while the on the right-hand side, where ω/B is large, an increase in K decreases ρ_q^{*2} (firms' attention to fundamental decreases with K).

C.8. First- and Second-Order Effects of Rational Inattention with Endogenous Capacity

Although our characterizations of the equilibrium capacities, κ, attention to the fundamental, ρ_q, and covariance of prices with the funddamentla, δ, in the model with endogenous choice of κ are analytical, they are only implicit. As a result, it is not possible to solve for κ or ρ_q as explicit functions of the parameters. In this section, I provide a second-order approximation of the equilibrium κ, ρ_q and δ in the model with endogenous choice of κ. These approximations are arbitrarily accurate when the cost of capacity approaches zero (i.e. $\omega/B=0$). By using these approximations, we can write κ and ρ_q as explicit functions of model parameters up to the first or second order, which provides further intuition for the comparative statics of the model.

Consider the unique equilibrium that arises in the model with endogenous capacity when $\omega < B(1-\alpha)^2$. Focusing on this equilibrium, we observe that it only arises when the ratio $\frac{\omega}{B}<(1-\alpha)^2<1$ is small. Moreover, recall that λ^* is strictly positive in this equilibrium and solves the following equations:

$$\lambda^*=1-\frac{\omega}{BV^*} \tag{C.25}$$

$$V^*=(\frac{1-\alpha}{1-\alpha\lambda^*})^2\frac{K-1+\alpha\lambda^*}{K-1+\alpha\lambda^*\frac{1-\alpha}{1-\alpha\lambda^*}} \tag{C.26}$$

which define λ^* as an implicit function of the parameters ω/B, α and K, $\lambda^*=\lambda^*(\omega/B,K,\alpha)$ (note that ω/B was originally $\omega/(B\times\mathbb{V}\text{ar}(q))$ where we have set $\mathbb{V}\text{ar}(q)=1$. So a small ω/B should be interpreted as a case where ω is small relative to B and/or the unconditional variance of q). Moreover, the set of ω/B's that satisfy the uniqueness condition also includes $\omega=0$ which corresponds to the frictionless benchmark with rational expectations and full information. It is straightforward to see that when $\omega\downarrow 0$, optimal capacity grows unboundedly towards infinity and $\lambda^*(\omega/B,\alpha,K)\uparrow 1$. In fact, compactifying $\mathbb{R}_+$ with the addition of $+\infty$, we can define this limit as the solution to the capacity choice problem of firms, where with $\omega=0$, $\kappa^*=+\infty$ and $\lambda^*=1$.

Moreover, once we have the equilibrium λ^*, we can calculate the attention of firms to the fundamental $\rho_q^{*2}=\rho_q^{*2}(\omega/B,\alpha,K)=\frac{K-1+\alpha\frac{1-\alpha}{1-\alpha\lambda^*}\lambda^*}{K-1+\alpha\lambda^*}$ as well as the comovement of prices with the fundamental $\delta^*=\delta^*(\omega/B,\alpha,K)=\frac{(1-\alpha)\lambda^*}{1-\alpha\lambda^*}$ also as functions of ω/B, α and K. Again, the case of $\omega=0$ provides a natural benchmark for these values as well, since with no cost of attention, firms pay full attention to the fundamental q, $\rho_q^{*2}(0,\alpha,K)=1$ and prices comove one to one with the fundamental $\delta(0,\alpha,K)=1$.

Since the unique equilibrium considered here arises for small values of $\omega/B<(1-\alpha)^2$, it is appropriate and useful to consider a second-order approximation of the equilibrium functions, $\lambda^*(\omega/B,\alpha,K)$, $\rho_q^{*2}(\omega/B,\alpha,K)$ and $\delta(\omega/B,\alpha,K)$ around the benchmark $\omega/B=0$. We derive these approximations below.

Approximation of $\lambda^*(\omega/B,\alpha,K)$. Note that when with $\omega/B<(1-\alpha)^2$, the implicit function defining $\lambda^*(\omega/B,\alpha,K)$ is smooth and continuously differentiable. A second order Taylor expansion is given by:

$$\lambda^*(\frac{\omega}{B},\alpha,K)=\overbrace{\lambda^*(0,\alpha,K)}^{=1}+[\frac{\partial\lambda^*}{\partial\frac{\omega}{B}}]_{\omega=0}\times\frac{\omega}{B}+\frac{1}{2}[\frac{\partial^2\lambda^*}{\partial\frac{\omega}{B}^2}]_{\omega=0}\times(\frac{\omega}{B})^2+\mathcal{O}(\|\frac{\omega}{B}\|^3) \tag{C.27}$$

Now, to calculate the first and second-order derivatives of the λ^* with respect to ω/B note that from $\lambda^*=1-\frac{\omega}{BV^*}$, we have:

$$\frac{\partial \lambda^*}{\partial \frac{\omega}{B}} = -\frac{1}{V^*}+\frac{\omega}{BV^{*2}}\frac{\partial V^*}{\partial \frac{\omega}{B}} \tag{C.28}$$

$$\frac{\partial^2 \lambda^*}{\partial \frac{\omega}{B}^2} = \frac{2}{V^{*2}}\frac{\partial V^*}{\partial \frac{\omega}{B}}+\frac{\omega}{B}\frac{\partial}{\partial \frac{\omega}{B}}\left(\frac{1}{V^{*2}}\frac{\partial V^*}{\partial \frac{\omega}{B}}\right) \tag{C.29}$$

Evaluating these at $\omega=0$ and noting that the first and second derivatives of V^* with respect to ω/B are finite so that multiplied by $\omega=0$ they are zero as well, we have:

$$[\frac{\partial \lambda^*}{\partial \frac{\omega}{B}}]_{\omega=0} = -[\frac{1}{V^*(\frac{\omega}{B},\alpha,K)}]_{\omega=0} = -1 \tag{C.30}$$

$$[\frac{\partial^2 \lambda^*}{\partial \frac{\omega}{B}^2}]_{\omega=0} = [\frac{2}{V^{*2}}\frac{\partial V^*}{\partial \frac{\omega}{B}}]_{\omega=0} = 2[\frac{\partial V^*(\frac{\omega}{B},\alpha,K)}{\partial \frac{\omega}{B}}]_{\omega=0} \tag{C.31}$$

where we have evaluated the value of $V^*(0,\alpha,K)$ using the equation for V^* above (Equation (9)). Also, note that for the seoncd derivative of λ^*, we need to calculate the first derivative of V^* with respect to ω/B which, by Equation (9), is only a function of ω/B through λ^*. Using the chain rule this derivative is given by:

$$\begin{aligned}\frac{\partial V^*}{\partial \frac{\omega}{B}} &= \frac{\partial V^*}{\partial \lambda^*}\frac{\partial \lambda^*(\frac{\omega}{B},\alpha,K)}{\partial \frac{\omega}{B}} \\ &= \underbrace{[2\frac{\alpha}{1-\alpha\lambda^*}+\frac{\alpha}{K-1+\alpha\lambda^*}-\frac{\alpha}{(1-\alpha\lambda^*)(K-1)+\alpha(1-\alpha)\lambda^*}\frac{1-\alpha}{1-\alpha\lambda^*}]V^*}_{=\frac{\partial V^*}{\partial \lambda^*}}\times\underbrace{(-\frac{1}{V^*}+\frac{\omega}{BV^{*2}}\frac{\partial V^*}{\partial \frac{\omega}{B}})}_{=\frac{\partial \lambda^*(\frac{\omega}{B},\alpha,K)}{\partial \frac{\omega}{B}}}\end{aligned} \tag{C.32}$$

Evaluating this at $\omega=0$ gives:

$$[\frac{\partial V^*}{\partial \frac{\omega}{B}}]_{\omega=0} = -\frac{\alpha}{1-\alpha}(1+\frac{K-1}{K-1+\alpha}) \tag{C.33}$$

Thus, plugging in the values of the derivatives to Equation (C.27), we have:

$$\lambda^*(\frac{\omega}{B},\alpha,K)=1-\frac{\omega}{B}-\frac{\alpha}{1-\alpha}(1+\frac{K-1}{K-1+\alpha})(\frac{\omega}{B})^2+\mathcal{O}(\|\frac{\omega}{B}\|^3) \tag{C.34}$$

Approximation of $\rho_q^{*2}(\omega/B,\alpha,K)$. Using Equations (C.6) and (C.7) in Appendix C.2, the correlation between the firm's price and the fundamental is given by

$$\rho_q^{*2} = \frac{K-1+\alpha\frac{(1-\alpha)\lambda^*}{1-\alpha\lambda^*}}{K-1+\alpha\lambda^*}\lambda^* \tag{C.35}$$

Thus, we see that ρ_q^{*2} is a function of α, K and ω/B through λ^*. A second order Taylor expansion of ρ_q^{*2} around $\omega=0$ is:

$$\rho_q^{*2}(\frac{\omega}{B},\alpha,K)=\overbrace{\rho_q^{*2}(0,\alpha,K)}^{=1}+[\frac{\partial \rho_q^{*2}}{\partial \frac{\omega}{B}}]_{\omega=0}\times\frac{\omega}{B}+\frac{1}{2}[\frac{\partial^2 \rho_q^{*2}}{\partial \frac{\omega}{B}^2}]_{\omega=0}\times(\frac{\omega}{B})^2+\mathcal{O}(\|\frac{\omega}{B}\|^3) \tag{C.36}$$

Now, since ρ_q^{*2} only depends on ω/B through λ^*, we can use the chain rule to calculate its first and second derivative as:

$$\frac{\partial\rho_q^{*2}}{\partial\frac{\omega}{B}} = \frac{\partial\rho_q^{*2}}{\partial\lambda^*}\frac{\partial\lambda^*}{\partial\frac{\omega}{B}} \tag{C.37}$$

$$\frac{\partial^2\rho_q^{*2}}{\partial\frac{\omega}{B}^2} = \frac{\partial^2\rho_q^{*2}}{\partial\lambda^{*2}}(\frac{\partial\lambda^*}{\partial\frac{\omega}{B}})^2 + \frac{\partial\rho_q^{*2}}{\partial\lambda^*}\frac{\partial^2\lambda^*}{\partial\frac{\omega}{B}^2} \tag{C.38}$$

While these are straightforward to derive, the expression for the second derivative is quite long. However, up to first-order, we have:

$$\rho_q^{2^*}(\frac{\omega}{B},\alpha,K) = 1-(1-\alpha\frac{K-1}{K-1+\alpha})\frac{\omega}{B(1-\alpha)} + \mathcal{O}(\|\frac{\omega}{B}\|^2) \tag{C.39}$$

Approximation of $\delta^*(\omega/B,\alpha,K)$**.** Recall from Equation (C.6) that the covariance of the aggregate price with the fundamental, denoted by δ^* is given by:

$$\delta^*(\frac{\omega}{B},\alpha,K) = \frac{(1-\alpha)\lambda^*}{1-\alpha\lambda^*} \tag{C.40}$$

which depends on ω/B and K through λ^*, but depends on α both directly as well as indirectly through λ^*. A second order Taylor expansion of δ^* around $\omega = 0$ is:

$$\delta^*(\frac{\omega}{B},\alpha,K) = \overbrace{\delta^*(0,\alpha,K)}^{=1} + [\frac{\partial\delta^*}{\partial\frac{\omega}{B}}]_{\omega=0} \times \frac{\omega}{B} + \frac{1}{2}[\frac{\partial^2\delta^*}{\partial\frac{\omega}{B}^2}]_{\omega=0} \times (\frac{\omega}{B})^2 + \mathcal{O}(\|\frac{\omega}{B}\|^3) \tag{C.41}$$

Now, since δ^* only depends on ω/B through λ^*, we can use the chain rule to calculate its first and second derivative as:

$$\frac{\partial\delta^*}{\partial\frac{\omega}{B}} = \frac{\partial\delta^*}{\partial\lambda^*}\frac{\partial\lambda^*}{\partial\frac{\omega}{B}} \tag{C.42}$$

$$\frac{\partial^2\delta^*}{\partial\frac{\omega}{B}^2} = \frac{\partial^2\delta^*}{\partial\lambda^{*2}}(\frac{\partial\lambda^*}{\partial\frac{\omega}{B}})^2 + \frac{\partial\delta^*}{\partial\lambda^*}\frac{\partial^2\lambda^*}{\partial\frac{\omega}{B}^2} \tag{C.43}$$

Noting that

$$\frac{\partial\delta^*}{\partial\lambda^*} = \frac{1-\alpha}{(1-\alpha\lambda^*)^2} \tag{C.44}$$

$$\frac{\partial^2\delta^*}{\partial\lambda^{*2}} = \frac{2\alpha(1-\alpha)}{(1-\alpha\lambda^*)^3} \tag{C.45}$$

Evaluating these at $\omega = 0$ and using Equations (C.30) and (C.31) we have

$$[\frac{\partial\delta^*}{\partial\frac{\omega}{B}}]_{\omega=0} = -\frac{1}{1-\alpha} \tag{C.46}$$

$$[\frac{\partial^2\delta^*}{\partial\frac{\omega}{B}^2}]_{\omega=0} = -2\frac{\alpha}{(1-\alpha)^2}(\frac{K-1}{K-1+\alpha}) \tag{C.47}$$

So that

$$\delta^*(\frac{\omega}{B},\alpha,K) = 1-\frac{\omega}{B(1-\alpha)} - \frac{(K-1)\alpha}{K-1+\alpha}(\frac{\omega}{B(1-\alpha)})^2 + \mathcal{O}(\|\frac{\omega}{B}\|^3) \tag{C.48}$$

C.9. Proofs of Propositions for the Static Model

This section includes the proofs of Propositions in the static model. The proofs and derivations for Section 4 are included in Appendix H.

Proof of Proposition 1

1. First, observe from Equations (C.6) and (C.7) in Appendix C.2 that the correlation between the firm's price and the fundamental is given by
$$\rho_q^{*2} = \frac{\mathbb{C}\text{ov}(p_{j,k},q)^2}{\mathbb{V}\text{ar}(p_{j,k})} = \frac{K-1+\alpha\delta}{K-1+\alpha\lambda}\lambda. \tag{C.49}$$
Moreover, notice that $\delta = \frac{1-\alpha}{1-\alpha\lambda}\lambda < \lambda$ as long as $\lambda > 0$ and $\alpha > 0$. This implies directly that $\rho_q^{*2} < \lambda$. Thus, the correlation between the firm's price and the mistakes of its competitors is strictly positive: $\rho_v^{*2} = \lambda - \rho_q^{*2} > 0$, meaning that firms pay attention to the mistakes of their competitors.
2. Shown in the proof of Lemma C.4.
3. Differentiating the correlation ρ_q^* with respect to the number of competitors K, we have
$$\frac{\partial \rho_q^{*2}}{\partial K}\frac{1}{\rho_q^{*2}} = \frac{\alpha(\lambda-\delta)}{(K-1+\alpha\lambda)(K-1+\alpha\delta)} > 0$$
Also, with respect to α:
$$\frac{\partial \rho_q^{*2}}{\partial \alpha}\frac{1}{\rho_q^{*2}} = \frac{(K-1)(\delta-\lambda)+(K-1+\alpha\lambda)\alpha\frac{\partial\delta}{\partial\alpha}}{(K-1+\alpha\delta)(K-1+\alpha\lambda)} < 0.$$
The inequality comes from $\delta - \lambda < 0$ and $\frac{\partial \delta}{\partial \alpha} = \delta\frac{\lambda-1}{(1-\alpha)(1-\alpha\lambda)} < 0$.

Proof of Proposition 2

First, observe that the aggregate price is given by
$$p \equiv J^{-1}K^{-1}\sum_{j,k\in J\times K} p_{j,k} = \delta q + \frac{1}{JK}\sum_{j,k\in J\times K} v_{j,k}$$
Since J is large and $v_{j,k}$'s are independent across industries, this average mistake across all the firms in the economy converges to zero by the law of large numbers as $J\to\infty$. Therefore, $p=\delta q$. Moreover, $\mathbb{E}^{j,k}[p_{j,-k}] = \frac{\mathbb{C}\text{ov}(s_{j,k},p_{j,-k})}{\mathbb{V}\text{ar}(p_{j,k})} s_{j,k} = \tilde{\lambda} p_{j,k}$ and $\mathbb{E}^{j,k}[p] = \frac{\mathbb{C}\text{ov}(s_{j,k},p)}{\mathbb{V}\text{ar}(p_{j,k})} p_{j,k} = \frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda}\lambda p_{j,k}$ where $\tilde{\lambda} = \frac{\lambda(K-1)+\alpha\lambda}{K-1+\alpha\lambda} > \lambda$ is defined as in the proof of Lemma C.4. So, $\overline{\mathbb{E}^{j,k}[p_{j,-k}]} = \tilde{\lambda}p$, $\overline{\mathbb{E}^{j,k}[p]} = \frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda}\lambda p$. Therefore,
$$\mathbb{C}\text{ov}(\overline{\mathbb{E}^{j,k}[p_{j,-k}]},p) \quad = \quad \tilde{\lambda}\mathbb{V}\text{ar}(p) > \frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda}\lambda\mathbb{V}\text{ar}(p) = \mathbb{C}\text{ov}(\overline{\mathbb{E}^{j,k}[p]},p).$$
Also, if $K\to\infty$ then $\tilde{\lambda}\to\lambda$ and $\mathbb{C}\text{ov}(\overline{\mathbb{E}^{j,k}[p]},p) \to \mathbb{C}\text{ov}(\overline{\mathbb{E}^{j,k}[p_{j,-k}]},p)$.

Now, note that conditional on realization of the aggregate price $|p - \overline{\mathbb{E}^{j,k}[p]}| = (1 - \frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda}\lambda)|p| > (1-\tilde{\lambda})|p| = |p - \overline{\mathbb{E}^{j,k}[p_{j,-k}]}|$.

Proof of Proposition 3

Recall that

$$p_{j,k} = (1-\alpha)\mathbb{E}_{j,k}[q] + \alpha\mathbb{E}_{j,k}[p_{j,-k}]$$

given the equilibrium strategy of other firms with decomposition $p_{j,-k} = \delta q + v_{j,-k}$ and using the fact that the equilibrium is a recommendation strategy (so that $S_{j,k} = p_{j,k}$), we have

$$\begin{aligned}\mathbb{E}_{j,k}[q] = E[q|p_{j,k}] &= \frac{\mathbb{C}\text{ov}(q,p_{j,k})}{\mathbb{V}\text{ar}(p_{j,k})} p_{j,k}\\ \mathbb{E}_{j,k}[p_{j,-k}] &= \delta E[q|p_{j,k}] + E[v_{j,-k}|p_{j,k}]\\ &= \delta \frac{\mathbb{C}\text{ov}(q,p_{j,k})}{\mathbb{V}\text{ar}(p_{j,k})} p_{j,k} + \frac{\mathbb{C}\text{ov}(v_{j,-k},p_{j,k})}{\mathbb{V}\text{ar}(p_{j,k})} p_{j,k}\end{aligned}$$

Now, let $\hat{\delta} \equiv \mathbb{C}\text{ov}(q,p_{j,k})$ and $\sigma_v \equiv \mathbb{V}\text{ar}(v_{j,-k})$. Note that we can re-write the equations above in terms of firms' optimal correlation choices ρ_q and ρ_v

$$\begin{aligned}\mathbb{E}_{j,k}[q] = E[q|p_{j,k}] &= \frac{1}{\hat{\delta}} \overbrace{\frac{\mathbb{C}\text{ov}(q,p_{j,k})^2}{\mathbb{V}\text{ar}(p_{j,k})\mathbb{V}\text{ar}(q)}}^{=\rho_q^2} p_{j,k}\\ \mathbb{E}_{j,k}[p_{j,-k}] &= \delta E[q|p_{j,k}] + E[v_{j,-k}|p_{j,k}]\\ &= \frac{\delta}{\hat{\delta}} \underbrace{\frac{\mathbb{C}\text{ov}(q,p_{j,k})^2}{\mathbb{V}\text{ar}(p_{j,k})\mathbb{V}\text{ar}(q)}}_{\rho_q^2} p_{j,k} + \frac{\sigma_v}{\hat{\delta}} \underbrace{\frac{\mathbb{C}\text{ov}(v_{j,-k},p_{j,k})}{\sqrt{\mathbb{V}\text{ar}(v_{j,-k})\mathbb{V}\text{ar}(p_{j,k})}}}_{\rho_v} \times \underbrace{\frac{\mathbb{C}\text{ov}(q,p_{j,k})}{\sqrt{\mathbb{V}\text{ar}(q)\mathbb{V}\text{ar}(p_{j,k})}}}_{\rho_q} p_{j,k}\end{aligned}$$

where we have used $\mathbb{V}\text{ar}(q) = 1$. Using the fact that in the equilibrium $\hat{\delta} = \delta \Leftrightarrow \mathbb{C}\text{ov}(q, p_{j,-k}) = \mathbb{C}\text{ov}(q,p_{j,k})$, and aggregating the above equations across all firms, we have

$$\begin{aligned}\overline{\mathbb{E}_{j,k}[q]} &= \frac{\rho_q^2}{\delta} p = \rho_q^2 q\\ \overline{\mathbb{E}_{j,k}[p_{j,-k}]} &= \rho_q^2 p + \rho_q \rho_v \sigma_v q\end{aligned}$$

where we have also used $p = \delta q$. Finally, note that from the first-order conditions of the firms' problem with respect to ρ_q and ρ_v, we have

$$\alpha \rho_q \sigma_v = (1-\alpha+\alpha\delta)\rho_v \quad s.t. \quad \rho_q^2 + \rho_v^2 = \lambda$$

Substituting these above we get

$$\begin{aligned}\overline{\mathbb{E}_{j,k}[p_{j,-k}]} &= \rho_q^2 \delta q + \alpha^{-1}(1-\alpha+\alpha\delta)\rho_v^2 q\\ &= \rho_q^2 \delta q + \alpha^{-1}(1-\alpha+\alpha\delta)(\lambda - \rho_q^2) q\\ &= \alpha^{-1}\delta q - (\alpha^{-1}-1)\rho_q^2 q\end{aligned}$$

Thus, to examine how these covariances change with the number of competitors, we have:

$$\begin{aligned}\partial_K \mathbb{C}\text{ov}(q, \overline{\mathbb{E}_{j,k}[q]}) &= \partial_K \rho_q^2 > 0\\ \partial_K \mathbb{C}\text{ov}(q, \overline{\mathbb{E}_{j,k}[p_{j,-k}]}) &= \partial_K \left[\alpha^{-1}\delta - (\alpha^{-1}-1)\rho_q^2\right] = -(\alpha^{-1}-1)\partial_K \rho_q^2 < 0\end{aligned}$$

These two equations show that as the number of competitors increases, aggregate prices comove more with firms' average expectations of q and comove less with their expectations of their competitors' prices. The reason is that, because we are keeping κ fixed, higher attention to the fundamental comes at the expense to lower attention to competitors' prices. Note also that in this step, we have already imposed $\partial_K \delta = 0$ because we have already derived the expression for δ and shown its independence from K in Equation (C.6) and proof of Lemma C.4, independent of the argument put forth here.

Now, to see how these two forces balance, note that

$$\begin{aligned}
p &= (1-\alpha)\overline{\mathbb{E}_{j,k}[q]} + \alpha \overline{\mathbb{E}_{j,k}[p_{j,-k}]} \\
\Rightarrow \mathbb{C}\text{ov}(p,q) &= (1-\alpha)\mathbb{C}\text{ov}(q,\overline{\mathbb{E}_{j,k}[q]}) + \alpha \mathbb{C}\text{ov}(q,\overline{\mathbb{E}_{j,k}[p_{j,-k}]}) \\
\Rightarrow \partial_K \mathbb{C}\text{ov}(p,q) &= (1-\alpha)\underbrace{\partial_K \mathbb{C}\text{ov}(q,\overline{\mathbb{E}_{j,k}[q]})}_{=\partial_K \rho_q^2} + \alpha \underbrace{\partial_K \mathbb{C}\text{ov}(q,\overline{\mathbb{E}_{j,k}[p_{j,-k}]})}_{=-(\alpha^{-1}-1)\partial_K \rho_q^2} \\
&= (1-\alpha)\partial_K \rho_q^2 - \alpha(\alpha^{-1}-1)\partial_K \rho_q^2 \\
&= 0
\end{aligned}$$

Therefore, $\delta = \mathbb{C}\text{ov}(p,q)$ does not change with K because higher attention to q is exactly offset with lower attention to competitors' prices with a fixed κ.

Proof of Proposition 4

Recall that a symmetric equilibrium is characterized by a λ^* that solves Equations (9) and (10):

$$\lambda^* = \max\{0, 1 - \frac{\omega}{BV^*}\} \tag{C.50}$$

$$V^* = (\frac{1-\alpha}{1-\alpha\lambda^*})^2 \frac{K-1+\alpha\lambda^*}{K-1+\alpha\lambda^* \frac{1-\alpha}{1-\alpha\lambda^*}} \tag{C.51}$$

The first step is to show that if $\omega < B(1-\alpha)^2$ then $\lambda^* = 0$ cannot be an equilibrium. To see this, suppose $\lambda^* = 0$. Then, the second equation implies that $V^* = (1-\alpha)^2$. Now, plugging this into the first equation we get

$$\lambda^* = \max\{0, 1 - \frac{\omega}{B(1-\alpha)^2}\} = 1 - \frac{\omega}{B(1-\alpha)^2} > 0 \tag{C.52}$$

which contradicts the assumption that $\lambda^* = 0$. Therefore, $\lambda^* = 0$ cannot be an equilibrium if $\omega < B(1-\alpha)^2$ and we can assume without loss of generality that $\lambda^* > 0$ and that it solves:

$$\lambda^* = 1 - \frac{\omega}{BV^*} \tag{C.53}$$

$$V^* = (\frac{1-\alpha}{1-\alpha\lambda^*})^2 \frac{K-1+\alpha\lambda^*}{K-1+\alpha\lambda^* \frac{1-\alpha}{1-\alpha\lambda^*}} \tag{C.54}$$

Plugging V^* into the first equation, we get:

$$1 - \lambda^* = \frac{\omega}{B(1-\alpha)^2}(1-\alpha\lambda^*)(1-\alpha+\alpha\frac{(1-\lambda^*)(K-1)}{K-1+\alpha\lambda^*}) \tag{C.55}$$

Now note that the left-hand side of the equation above is strictly decreasing in λ^* and ranges from 1 to 0 as λ^* goes from 0 to 1. The right-hand side is also strictly decreasing in λ^* and goes from $\frac{\omega}{B(1-\alpha)^2} < 1$

to $\omega/B>0$ as λ^* goes from 0 to 1. Since the range of the left-hand side is a strict subset of the range of the right-hand side for $\lambda^*\in[0,1]$ and both sides are strictly decreasing in λ^*, then there must be a unique $\lambda^*>0$ that solves this equation as long as $\omega<B(1-\alpha)^2$.

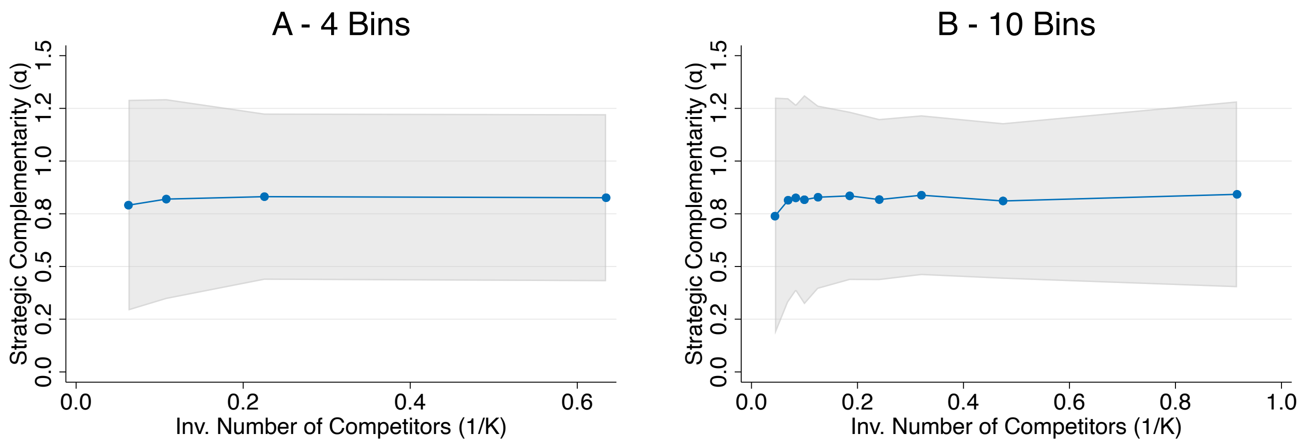
Now to see how λ^* varies with ω,B,α and K, note that larger values of ω, α or K shift the right-hand side upwards and move the intersection to the left, decreasing λ^*. A larger B, on the other hand, shifts the right-hand side downwards and moves the intersection to the right, increasing λ^*. Therefore, λ^* is decreasing in ω,α and K and increasing in B.

D Strategic Complementarity and Number of Competitors in Survey Data

As discussed in Section 4.2, micro-foundations of strategic complementarity relate this object to firms' market shares in the equilibrium. Since around the symmetric steady-state market share is the inverse of the number of firms in the oligopoly, these micro-foundations relate strategic complementarity to the number of competitors. Moreover, we have shown that different micro-foundations have different implications for this relationship. For instance, the two-layer CES structure, as in Atkeson and Burstein (2008), on its own implies that strategic complementarity should increase with market share and thus decrease with the number of competitors (for example, recall that in the benchmark model of this paper, without decreasing returns to scale, $\alpha=\frac{1-\eta^{-1}}{K}$). However, more general aggregators might reverse this relationship (see, e.g., Wang and Werning (2022)'s discussion of this relationship in the Kimball model, which is also derived in Appendix G in this paper as $\alpha=\frac{\zeta(K-2)+(1-\eta^{-1})^2}{\zeta(K-2)+(1-\eta^{-1})K}$, where $\zeta=0$ nests CES but in general is related to superelasticity of demand. For $\zeta>1$ we can see that α increases with K similar to Wang and Werning (2022)). Moreover, the decreasing returns to scale also add a force that makes strategic complementarity increase with the number of competitors, as is the case in the calibrated model of Section 4. While different models have different predictions for how strategic complementarity varies with the number of competitors, we can examine this relationship empirically in the survey data.

One issue that needs to be addressed is that variation in K in the survey is not that large which makes the estimates of α noisier for larger K. To address this, I divide the data into equally sized bins, in terms of the number of observations, as a function of the inverse of the number of competitors (which corresponds to steady-state market share in the model). Panels A and B in Figure D.1 present connected lines with average α by $1/K$ bin, after controlling for industry dummies. Shaded areas denote 1 standard deviation.

Figure D.1: Strategic Complementarity as a Function of $1/K$ (± 1 Standard Deviation)



Notes: This figure presents a binned plot where dots denote average α over equally sized bins of the inverse of the number of competitors $1/K$ after controlling for industry fixed effects. The shaded area denotes $+/-1$ standard deviation.

We can make the following observations from this graph. First, quantitatively, α is somewhat flat across $1/K$ bins, with a slight increase around the first quartiles in Panel A, which is qualitatively consistent with studies that document strategic complementarity to decline with market share (Amiti, Itskhoki, & Konings, 2019). Second, repeating the exercise with deciles of $1/K$ in Panel B, α seems to increase slightly and then decrease hinting at a slight non-monotonicity. However, the magnitude of changes based on the point estimates seems to be relatively small, ranging from 0.8 to slightly below 0.9. To make this observation more rigorously, Table D.1 regresses strategic complementarity on different quartiles of $1/K$ and while hinting at the slight non-monotonicity seen in Figure D.1, it shows that we cannot reject the null-hypothesis that strategic complementarity is not different across different quartiles.

	Dep. Variable: Strategic Complementarity α			
	(1)		(2)	
Constant	0.827***	(0.018)	0.805***	(0.022)
First Quartile $1/K$	0.039	(0.046)	0.038	(0.046)
Second Quartile $1/K$	0.099**	(0.050)	0.100**	(0.050)
Third Quartile $1/K$	0.013	(0.054)	0.011	(0.055)
Fourth Quartile $1/K$	-	-	-	-
Observations	2,824		2,823	
Industry dummies	No		Yes	

Table D.1: Differential Strategic Complementarity by $1/K$ quartiles

Notes: This table estimates how strategic complementarity varies across different quartiles of $1/K$ relative to its fourth quartile. Robust standard errors in parenthesis. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

E Available Information in the Dynamic Model

The set of available signals in the dynamic model is an extension of the set defined in Appendix C.1. The key notion in this extension is that nature draws new shocks every period, and the set of available information in the economy expands to incorporate these new realizations. To capture this evolution, I define a signal structure as a sequence of sets $(\mathbb{S}^t)_{t=-\infty}^{\infty}$ where $\mathbb{S}^{t-s} \subset \mathbb{S}^t, \forall s \geq 0$. Here, $\mathbb{S}^t$ denotes the set of available signals at time t, and it contains all the previous sets of signals that were available in previous periods.

To construct the signal structure, suppose that every period, in addition to the shock to the nominal demand, the nature draws countably infinite uncorrelated standard normal noises. Similar to Appendix C.1, let $\mathbb{S}_t$ be the set of all finite linear combinations of these uncorrelated noises along with the newest innovation to q_t. Now, define $\mathbb{S}^t = \{\sum_{s=0}^{\infty} a_\tau e_{t-\tau} | \forall \tau \geq 0, a_\tau \in \mathbb{R}, e_{t-\tau} \in \mathbb{S}_{t-\tau}\}, \forall t$. First, for all t, $q_t \in \mathbb{S}^t$, as it is a linear combination of all $u_{t-\tau}$'s and $u_{t-\tau} \in \mathbb{S}_{t-\tau}, \forall \tau \geq 0$. This implies that while perfect information is available about the fundamentals of the economy, signals with arbitrarily less precision are also available for the firms, should they choose to acquire them.

F Derivations for the Dynamic Model

F.1. Solution to Household's Problem (16)

Let $\beta^t \varphi_{1,t}$ and $\beta^t \varphi_{2,t}$ be the Lagrange multipliers on household's budget and aggregation constraints, respectively. For ease of notation let $\mathcal{C}_{j,t} \equiv (C_{j,1,t}, \ldots, C_{j,K_j,t})$ be the vector of household's consumption from firms in industry $j \in J$, so that $C_{j,t} \equiv \Phi_j(\mathcal{C}_{j,t})$ where $\Phi_j(.)$ is an aggregator function that is homogenous of degree one and at least thrice differentiable in its arguments (note that this embeds the CES aggregator in the main text as well as the Kimball aggregator discussed in Appendix G). Moreover, for less crowded notation, I drop subscript j for ϕ_j and K_j whenever the industry index is implied from context. First, I derive the demand of the household for different goods. The first order condition with respect to $C_{j,k,t}$ is

$$P_{j,k,t} = \frac{1}{J} \frac{\varphi_{2,t}}{\varphi_{1,t}} C_t \frac{\Phi_k(\mathcal{C}_{j,t})}{\Phi(\mathcal{C}_{j,t})} \tag{F.1}$$

where $\Phi_k(\mathcal{C}_{j,t}) \equiv \frac{\partial \Phi(\mathcal{C}_{j,t})}{\partial C_{j,k,t}}$. Given these optimality conditions, we can show that total sales in the economy is proportional to aggregate output:

$$\sum_{(j,k) \in J \times K} P_{j,k,t} C_{j,k,t} = \frac{1}{J} \frac{\varphi_{2,t}}{\varphi_{1,t}} C_t \sum_{j \in J} \underbrace{\sum_{k \in K} \frac{\Phi_k(\mathcal{C}_{j,t})}{\Phi(\mathcal{C}_{j,t})} C_{j,k,t}}_{=1, \forall j \in J} = \frac{\varphi_{2,t}}{\varphi_{1,t}} C_t$$

where the equality under curly bracket is from Euler theorem for homogeneous function $\Phi(.)$. Therefore, $P_t \equiv \frac{\varphi_{2,t}}{\varphi_{1,t}}$ is the price of the aggregate consumption basket C_t and we can write $Q_t = P_t C_t$ as the nominal demand of the household for the aggregate consumption good. Now, for the particular case of the CES function in the main case, Equation (F.1) becomes:

$$P_{j,k,t} = (JK_j)^{-1} Q_t C_{j,k,t}^{-\eta^{-1}} C_{j,t}^{\eta^{-1}-1} \Rightarrow \sum_{k \in K_j} P_{j,k,t}^{1-\eta} = (JK_j)^{\eta-1} K_j Q_t^{1-\eta} C_{j,t}^{\eta-1} \tag{F.2}$$

where the right hand side follows from raising the left hand side to the power of $1-\eta$ and summing over k. Now, raising the right hand side to the power of $-\eta$ and dividing it by the left hand side gives the demand curve in the text:

$$C_{j,k,t}=Q_t\mathcal{D}(P_{j,k,t},P_{j,-k,t}), \quad \mathcal{D}(P_{j,k,t},P_{j,-k,t})\equiv\frac{1}{J}\frac{P_{j,k,t}^{-\eta}}{\sum_{k\in K_j}P_{j,k,t}^{1-\eta}} \tag{F.3}$$

Now, for a general Φ: from Equation (F.1), $\mathcal{P}_{j,t}\equiv(P_{j,1,t},...,P_{j,K,t})=\nabla\log(\Phi(\frac{\mathcal{C}_{j,t}}{J^{-1}P_tC_t}))$. I need to show that this function is invertible to prove that a demand function exists. For ease of notation, define function $f:\mathbb{R}^K\to\mathbb{R}^K$ such that $f(\mathbf{x})\equiv\nabla\log(\Phi(\mathbf{x}))$. Notice that $f(.)$ is homogeneous of degree -1, and the m,n'th element of its Jacobian, denoted by matrix $\mathcal{J}^f(\mathbf{x})$, is given by $\mathcal{J}^f_{m,n}(\mathbf{x})\equiv\frac{\partial}{\partial x_n}\frac{\Phi_m(\mathbf{x})}{\Phi(\mathbf{x})}=\frac{\Phi_{m,n}(\mathbf{x})}{\Phi(\mathbf{x})}-\frac{\Phi_n(\mathbf{x})}{\Phi(\mathbf{x})}\frac{\Phi_m(\mathbf{x})}{\Phi(\mathbf{x})}$. Let $\mathbf{1}$ be the unit vector in $\mathbb{R}^K$. Since $\Phi(.)$ is symmetric along its arguments, for any $k\in(1,...,K)$, $\Phi_1(\mathbf{1})=\Phi_k(\mathbf{1})$, $\Phi_{11}(\mathbf{1})=\Phi_{kk}(\mathbf{1})<0$. Since $\Phi(.)$ is homogeneous of degree 1, by Euler's theorem we have $\Phi(\mathbf{1})=\sum_{k\in K}\Phi_k(\mathbf{1})=K\Phi_1(\mathbf{1})$. Also, since $\Phi_k(.)$ is homogeneous of degree zero.[48] Similarly we have $0=0\times\Phi_k(\mathbf{1})=\sum_{l\in K}\Phi_{kl}(\mathbf{1})$. So, for any $l\neq k$, $\Phi_{kl}(\mathbf{1})=-\frac{1}{K-1}\Phi_{11}(\mathbf{1})>0$. This last equation implies that $\mathcal{J}^f(\mathbf{1})$ is an invertible matrix.[49] Therefore, by inverse function theorem $f(.)$ is invertible in an open neighborhood around $\mathbf{1}$, and therefore any symmetric point $\mathbf{x}=x.\mathbf{1}$ such that $x>1$. We can write $\frac{\mathcal{C}_{j,t}}{J^{-1}P_tC_t}=f^{-1}(\mathcal{P}_{j,t})$. It is straight forward to show that $f^{-1}(.)$ is homogeneous of degree -1 because $f(\mathbf{x})$ is homogeneous of degree -1: for any $\mathbf{x}\in\mathbb{R}^K$, $f^{-1}(a\mathbf{x})=f^{-1}(af(f^{-1}(\mathbf{x}))=f^{-1}(f(a^{-1}f^{-1}(\mathbf{x}))=a^{-1}f^{-1}(\mathbf{x})$. Now, $C_{j,k,t}=J^{-1}P_tC_tf_k^{-1}(\mathcal{P}_{j,t})$, where $f_k^{-1}(\mathbf{x})$ is the k'th element of the vector $f^{-1}(\mathcal{P}_{j,t})$. Finally, since $f(.)$ is symmetric across its arguments, so is $f^{-1}(\mathcal{P}_{j,t})$, meaning that $f_k^{-1}(\mathcal{P}_{j,t})=f_1^{-1}(\sigma_{k,1}(\mathcal{P}_{j,t}))$, where $\sigma_{k,1}(\mathcal{P}_{j,t})$ is a permutation that changes the places of the first and k'th element of the vector $\mathcal{P}_{j,t}$. Now, to get the notation in the text let $(P_{j,k,t},P_{j,-k,t})\equiv\sigma_{k,1}(\mathcal{P}_{j,t})$ and $\mathcal{D}(\mathbf{x})\equiv J^{-1}f_1^{-1}(\mathbf{x})$, which gives us the notation in the text:$C_{j,k,t}=P_tC_t\mathcal{D}(P_{j,k,t},P_{j,-k,t})$, where $\mathcal{D}(.,.)$ is homogeneous of degree -1. Finally, the optimality conditions of the household's problem with respect to B_t,C_t and L_t are straight forward and are given by $P_tC_t=\beta(1+i_t)\mathbb{E}_t^f[P_{t+1}C_{t+1}]$ and $P_tC_t=W_t$.

F.2. Quadratic Approximation to Firms' Profits

Define a firm's revenue net of its production costs at a given time as

$$\Pi(P_{j,k,t},P_{j,-k,t},Q_t)=P_{j,k,t}Q_t\mathcal{D}(P_{j,k,t},P_{j,-k,t})-(1-\bar{\mathrm{s}}_j)Q_t^{2+\gamma}\mathcal{D}(P_{j,k,t},P_{j,-k,t})^{1+\gamma} \tag{F.4}$$

Now for any given set of signals over time that firm j,k could choose to see, its profit maximization problem is

$$\max_{(P_{j,k,t}:S_{j,k}^t\to\mathbb{R})_{t=0}^\infty}\mathbb{E}[\sum_{t=0}^\infty\beta^tQ_t^{-1}\Pi(P_{j,k,t},P_{j,-k,t},Q_t)|S_{j,k}^{-1}]=\max_{(P_{j,k,t}:S_{j,k}^t\to\mathbb{R})_{t=0}^\infty}\mathbb{E}[\sum_{t=0}^\infty\beta^t\Pi(\frac{P_{j,k,t}}{Q_t},\frac{P_{j,-k,t}}{Q_t},1)|S_{j,k}^{-1}].$$

[48] Follows from homogeneity of $\Phi(\mathbf{x})$. Notice that $\Phi(a\mathbf{x})=a\Phi(\mathbf{x})$. Differentiate with respect to k'th argument to get $\Phi_k(a\mathbf{x})=\Phi_k(\mathbf{x})$.

[49] With some algebra, we can show that $\mathcal{J}^f(\mathbf{1})=\frac{\Phi_{11}(\mathbf{1})}{K-1}\mathbf{I}-\frac{\Phi_{11}(\mathbf{1})+K^{-1}}{K(K-1)}\mathbf{1}\mathbf{1}'$, meaning that $\mathcal{J}^f(\mathbf{1})$ is a symmetric matrix whose diagonal elements are strictly different than its off-diagonal elements. Hence, it is invertible.

where the second equality follows from the fact that the profit function is homogeneous of degree 1 as $\mathcal{D}(.,.)$ is homogeneous of degree -1. Now, let small letters denote logs of corresponding variables so that $p_{j,k,t}-q_t \equiv \ln(P_{j,k,t}/Q_t)$ and $p_{j,-k,t}-q_t \equiv \ln(P_{j,-k,t}/Q_t)$ and define the loss function of the firm from mispricing at a given time as

$$L(p_{j,k,t}-q_t,p_{j,-k,t}-q_t) \equiv \Pi(\frac{P^*_{j,k,t}}{Q_t},\frac{P_{j,-k,t}}{Q_t},1)-\Pi(\frac{P_{j,k,t}}{Q_t},\frac{P_{j,-k,t}}{Q_t},1),$$

where $P^*_{j,k,t}=\text{argmax}_x \Pi(x,P_{j,-k,t},Q_t)$ is the firms' optimal price for the particular realizations of Q_t and $P_{j,-k,t}$. Now note that

$$\min_{(p_{j,k,t}:S^t_{j,k}\to\mathbb{R})^\infty_{t=0}} \mathbb{E}[\sum_{t=0}^{\infty}\beta^t L(p_{j,k,t}-q_t,p_{j,-k,t}-q_t)|S^{-1}_{j,k}]$$

has the same solution as profit maximization problem of the firm. Moreover, recall from the main text that in the symmetric equilibrium of the full-information economy $\frac{P_{j,k,t}}{Q_t}=\frac{P_{j,-k,t}}{Q_t}=1$. Taking a second-order approximation to the net present value of firm's losses at a given time around the symmetric full-information equilibrium, we arrive at:

$$\sum_{t=0}^{\infty}\beta^t L(p_{j,k,t}-q_t,p_{j,-k,t}-q_t) \approx \underbrace{-\frac{1}{2}\Pi_{11}(1,1,1)}_{>0}\sum_{t=0}^{\infty}\beta^t(p_{j,k,t}-p^*_{j,k,t})^2,$$

where $p^*_{j,k,t}$ is such that $\Pi_1(\exp(p^*_{j,k,t})/Q_t,P_{j,-k,t}/Q_t,1)=0$, meaning that

$$p^*_{j,k,t}=q_t+\underbrace{\left(1+\frac{\Pi_{13}(1,1,1)}{\Pi_{11}(1,1,1)}\right)}_{\text{strategic complementarity}=\alpha_j}\times\frac{1}{K_j-1}\sum_{l\neq k}(p_{j,l,t}-q_t) \tag{F.5}$$

$$=(1-\alpha_j)q_t+\alpha_j p_{j,-k,t} \tag{F.6}$$

It is straightforward to calculate the derivatives $\Pi_{11}(1,1,1)$ and $\Pi_{13}(1,1,1)$ as

$$\Pi_{11}(1,1,1)=-\mathtt{rs}_j(\varepsilon^{\varepsilon}_{j,D}+(1+\gamma)(\varepsilon^j_D-1)) \tag{F.7}$$

$$\Pi_{13}(1,1,1)=\mathtt{rs}_j(\varepsilon^j_D-1) \tag{F.8}$$

where $\mathtt{rs}_j \equiv \mathcal{D}(1,1)=(JK_j)^{-1}$ is the revenue share (or relative size) of the firm in the symmetric full-information equilibrium, ε^j_D is the demand elasticity and $\varepsilon^{\varepsilon}_{j,D}$ is the superelasticity of demand for a firm in sector j in the full-information symmetric equilibrium. Note that this gives a general expression for strategic complementarity as:

$$\alpha_j=1-\frac{\varepsilon^j_D-1}{\varepsilon^{\varepsilon}_{j,D}+(1+\gamma)(\varepsilon^j_D-1)} \tag{F.9}$$

Thus, note that we can write Π_{11} as:

$$\Pi_{11}=-\mathtt{rs}_j\frac{\varepsilon^j_D-1}{1-\alpha_j} \tag{F.10}$$

and the firm's objective for its attention problem is therefore given by

$$\max_{\{\kappa_{j,k,t},S_{j,k,t},p_{j,k,t}(S^t_{j,k})\}_{t\geq 0}} -\mathtt{rs}_j \mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t\left(\underbrace{\frac{1}{2}\frac{\varepsilon^j_D-1}{1-\alpha_j}(p_{j,k,t}(S^t_{j,k})-p^*_{j,k,t})^2}_{\text{loss from mispricing}}+\underbrace{(1-\mathtt{s}_j)\omega\kappa_{j,k,t}}_{\text{cost of capacity}}|S^{-1}_{j,k}\right)\right] \tag{F.11}$$

Dividing the objective by $1-\mathtt{s}_j=\frac{\varepsilon^j_D-1}{\varepsilon^j_D}$ gives us the objective in the main text where

$$B_j=\frac{\varepsilon^j_D}{1-\alpha_j}=\frac{\eta-(\eta-1)K_j^{-1}}{(1-(1-\eta^{-1})K_j^{-1})(\frac{1+\gamma}{1+\gamma\eta(1-(1-\eta^{-1})K_j^{-1})^2})}=\frac{\eta+\gamma(\eta-(\eta-1)K_j^{-1})^2}{1+\gamma} \tag{F.12}$$

where the equalities follow from the expression for demand elasticities and strategic complementarities in the main text.

G Strategic Complementarity under Kimball Demand

In the paper's main text, I consider a nested CES aggregator and derive the strategic complementarities under the demand system implied by that aggregator. An alternative approach in the literature is using the Kimball aggregator but mainly used with monopolistic competition. In this section, I derive the demand functions of firms given this aggregator in an *oligopolistic* setting for comparison.

The Kimball aggregator assumes that the function $\Phi(C_{j,1,t},\ldots,C_{j,K,t})$ is implicitly defined by

$$1=K^{-1}\sum_{k\in K}f(\frac{KC_{j,k,t}}{\Phi(C_{j,1,t},\ldots,C_{j,K,t})}), \tag{G.1}$$

where $f(.)$ is at least thrice differentiable, and $f(1)=1$ (so that $\Phi(1,\ldots,1)=K$). Observe that this coincides with the CES aggregator when $f(x)=x^{\frac{\eta-1}{\eta}}$. To derive the demand functions, recall that the first order conditions of the household's problem are $P_{j,k,t}=J^{-1}Q_t\frac{\frac{\partial}{\partial C_{j,k,t}}C_{j,t}}{C_{j,t}},\forall j,k$where $C_{j,t}=\Phi(C_{j,1,t},\ldots,C_{j,K,t})$. Implicit differentiation of Equation (G.1) gives

$$P_{j,k,t}=J^{-1}Q_t\frac{f'(\frac{KC_{j,k,t}}{C_{j,t}})}{\sum_{l\in K}C_{j,l,t}f'(\frac{KC_{j,l,t}}{C_{j,t}})},\forall j,k. \tag{G.2}$$

To invert these functions and get the demand for every firm in terms of their competitors' prices, guess that there exists a function $F:\mathbb{R}^K\to\mathbb{R}$ such that $\frac{\sum_{l\in K}C_{j,l,t}f'(\frac{KC_{j,l,t}}{C_{j,t}})}{J^{-1}Q_t}=F(P_{j,1,t},\ldots,P_{j,K,t})$. I verify this guess by plugging in this guess to Equation (G.2), which implies the function $F(.)$ is implicitly defined by $1=K^{-1}\sum_{k\in K}f(f'^{-1}(P_{j,l,t}F(P_{j,1,t},\ldots,P_{j,K,t})))$. Note that this is consistent with the guess and $F(.)$ only depends on the vector of these prices. It is straight forward to show that $F(.)$ is symmetric across its arguments and homogeneous of degree -1.[50] Now, given these derivations, we can derive the demand function of firm j,k as a function of the aggregate demand, its own price and the prices of its competitors.

[50]Symmetry is obvious to show. To see homogeneity, differentiate the implicit function that defines $F(.)$ with respect to each of its arguments and sum up those equations to get that for any $X=(x_1,\ldots,x_K)\in\mathbb{R}^K$, $-F(X)=\sum_{k\in K}x_k\frac{\partial}{\partial x_k}F(X)$. Now, notice that for any $a\in\mathbb{R},X\in\mathbb{R}^K$, $\frac{\partial aF(aX)}{\partial a}=0$. Thus, for any $X\in\mathbb{R}^K$, $aF(aX)$ is independent of a, and in particular $aF(aX)=F(X)\Rightarrow F(aX)=a^{-1}F(X)$.

Similar to the main text we can write this as

$$C_{j,k,t} = J^{-1} Q_t D(P_{j,k,t}, P_{j,-k,t}),\, D(P_{j,k,t}, P_{j,-k,t}) \equiv \frac{f'^{-1}(P_{j,k,t} F(P_{j,1,t}, \ldots, P_{j,K,t}))}{\sum_{l \in K} P_{j,l,t} f'^{-1}(P_{j,l,t} F(P_{j,1,t}, \ldots, P_{j,K,t}))}$$

In the spirit of the CES aggregator I define $\eta \equiv -\frac{f'(1)}{f''(1)}$ as the inverse of the elasticity of $f'(x)$ at $x=1$, and assume $\eta > 1$. It is straightforward to show that η is the elasticity of substitution between industry goods around a symmetric point. Moreover, the elasticity of demand for every firm around a symmetric point is $\eta - (\eta - 1)K^{-1}$ similar to the case of a CES aggregator. Also, define $\zeta(x) \equiv \frac{\partial \log(-\frac{\partial \log(f'(x))}{\partial \log(x)})}{\partial \log(x)}$ as the elasticity of the elasticity of $f'(x)$: $\zeta(x) = \frac{f'''(x)}{f''(x)} x - \frac{f''(x)}{f'(x)} x + 1$. For notational ease let $\zeta \equiv \zeta(1)$ and assume $\zeta \geq 0$ ($\zeta = 0$ corresponds to the case of CES aggregator). These assumptions ($\eta > 1$ and $\zeta \geq 0$ are sufficient for weak strategic complementarity, $\alpha \in [0,1)$). While the usual approach in the literature is to assume $K \to \infty$ and look at super elasticities in this limit, a part of my main results revolve around the finiteness of the number of competitors and the fact that the degree of strategic complementarity is decreasing in K. Therefore, I derive the degree of strategic complementarity for any finite K. With some intense algebra we get $\alpha = \frac{\zeta(K-2) + (1-\eta^{-1})^2}{\zeta(K-2) + (1-\eta^{-1})K} \in [0,1)$. This imbeds the CES aggregator when $\zeta = 0$, in which case $\alpha = (1-\eta^{-1})K^{-1}$.

H Proofs of Propositions for the Dynamic Model

Proof of Proposition 5

This proof has two parts. Part I casts a firm's problem into the abstract problem studied in Afrouzi and Yang (2019) and then applies Lemmas 1 and 3 from that paper,[51] concluding that it is optimal for firms to always observe one Gaussian signal at any given time t, for any $\beta \in [0,1)$. Part II of the proof then derives the optimal shape of the signal under the assumption of $\beta = 0$ and shows that these signals take the form of "ideal price plus noise."

Part I (Optimality of One Signal at Each Time for $\beta \in [0,1)$). Let $(S_{l,m}^{-1})_{(l,m) \in J \times K}$ denote the initial signal structure of the economy that firms inherit at time 0. Pick any firm j,k as the firm whose problem is being studied here and, to economize on notation, drop (j,k) when it is clear from the context. Consider a strategy profile for all other firms in the economy, denoted by $\varsigma = (S_{l,m,t} \subset \mathbb{S}^t, p_{j,l,t} : S_{l,m}^t \to \mathbb{R})_{(l,m) \neq (j,k)}^{t \geq 0}$. Define $\vec{x}_t(\varsigma) \equiv (q_t, p_{l,m,t}(S_{l,m}^t))_{(l,m) \neq (j,k)}$ and $X^t(\varsigma) \equiv \{\vec{x}_{j,k,\tau}(\varsigma) : 0 \leq \tau \leq t\}$. Note that under strategy profile ς, $X^t(\varsigma)$ has a stochastic process that is exogenous to firm j,k and is taken as given by that firm. Moreover, note that $X^t(\varsigma)$ contains all the variables that firm j,k would potentially pay attention to at time t subject to the feasibility of available information, as captured by $\mathbb{S}^t$. Define also the function $v_j(.)$ as the firm j,k's losses from mispricing at time t under price p_t and strategy ς as

$$v_j(p_t, \vec{x}_t(\varsigma)) \equiv -\frac{1}{2} B_j (p_t - p^*(\vec{x}_t(\varsigma)))^2, \quad p^*(\vec{x}_t(\varsigma)) \equiv (1-\alpha_j) q_t + \alpha_j p_{j,-k,t}(\varsigma) \qquad \text{(H.1)}$$

[51] https://afrouzi.com/dynamic_inattention/draft_2019_10.pdf#page=7.

With slight abuse of notation, also let $S^{t-1} = S^{t-1}_{j,k}$, denote firm j,k's information set at time $t-1$. Then, we can re-write the problem of firm j,k in Equation (26) as

$$\max_{\{S_t \subset \mathbb{S}^t, p_t(S^t)\to\mathbb{R}\}_{t\geq 0}} \sum_{t=0}^{\infty} \beta^t \mathbb{E}[v_j(p_t;\vec{x}_t(\varsigma)) - \omega \mathcal{I}(X^t(\varsigma);S^t|S^{t-1})|S^{-1}] \tag{H.2}$$
$$\text{subject to} \quad S^t = S^{t-1} \cup \{S_t\}, \quad \forall t \geq 0, \quad S^{-1} \text{ given.}$$

where we have substituted the information processing constraint into the objective in Equation (26).[52] We can see that, for a ς that firm j,k takes a given, Equation (H.2) is exactly the problem studied in (Afrouzi & Yang, 2019, RI Problem, p. 7).[53] Applying Lemma 1 from that paper we conclude that if $\{S^t \subset \mathbb{S}^t, p_t : S^t \to \mathbb{R}\}$ is a solution to Equation (H.2), then—letting $p^t \equiv \{p_\tau : 0 \leq \tau \leq t\} \cup S^{-1}$—$X^t(\varsigma) \to p^t \to S^t$ forms a Markov chain for all $t \geq 0$; i.e.,

$$X^t(\varsigma) \perp S^t | p^t \iff \mathcal{I}(X^t(\varsigma);S^t|p^t) = 0 \tag{H.3}$$

Thus, under the optimal information structure and pricing strategies, p^t is a sufficient statistic for S^t concerning $X^t(\varsigma)$, which means that the firm's pricing history reveals all of its acuqired information up to time t. Moreover, since signals are Gaussian (by Lemma 3 in Afrouzi and Yang (2019)), p^t is also a Gaussian process and $p^t \in \mathbb{S}^t$. Thus, p^t is a recommendation strategy for firm i that weakly dominates S^t. Therefore, recommendation strategies are also optimal for the dynamic problem, and firms prefer to observe one signal per period of time which recommends the price they should charge in that period.

Part II (Shape of the Optimal Signals with $\beta = 0$). While we have shown that with any $\beta \in [0,1)$ it is optimal for firms to observe only one signal at any given time, we have not characterized the shape of this optimal signal, particularly how it loads on different shocks. In general, optimal signals can have complicated representations, but the case of $\beta = 0$ is special. We can show that in this case, optimal signals are of the intuitive form as in the static model of "ideal price plus noise." This is, however, not ex ante obvious, since in contrast to the static model, shocks in the dynamics model can be auto-correlated and firms can choose to pay attention to past realizations of the state vector, $\vec{x}_{t-\tau}(\varsigma) : \tau \geq 1$. If we make the additional assumption that $\vec{x}_t(\varsigma)$ is a Markov process, we can directly apply the first-order conditions from Afrouzi and Yang (2019) to prove that optimal signals take the shape above. But, for $\beta = 0$, we can characterize this result more generally without the Markov assumption, which is different here from Afrouzi and Yang (2019). The rest of this part is devoted to proving this result. (Later, to solve the dynamic model numerically, we approximate $\vec{x}_t(\varsigma)$ with a Markov process and use the numerical methods from Afrouzi and Yang (2019) to solve for the shape of optimal signals for a calibrated value of $\beta > 0$, which no longer take the form ideal price plus noise).

Similar to Part I, fix a firm $(j,k) \in J \times K$ and consider its problem at time t for a given strategy of

[52] Here we are using the result that the information processing capacity in Equation (26) always binds. To see why, suppose that the constraint does not bind for some t. Then, the firm produces information capacity that is not used for acquiring information. Thus, the firm would be strictly better off reducing the production capacity for some small ϵ without affecting its information structure, implying that the constraint should always bind at the optimum.

[53] See https://afrouzi.com/dynamic_inattention/draft_2019_10.pdf#page=7.

other firms in the economy, which we denote by ς. To economize on notation, we continue dropping (j,k) when it is clear from the context. Let $\vec{x}_t(\varsigma)$, $X^t(\varsigma)$, and $p^*(x_t(\varsigma))$ defined as in Part I. Note that the firm's ideal price, $p^*(x_t(\varsigma))$, can be written as:

$$p^*(\vec{x}_t(\varsigma)) = \mathbf{w}'\vec{x}_t, \quad \mathbf{w}' \equiv (1-\alpha_j, \underbrace{\frac{\alpha_j}{K_j-1}, \ldots, \frac{\alpha_j}{K_j-1}}_{K_j-1 \text{ times}}, \underbrace{0,0,\ldots,0}_{(J-1)\times K_j \text{ times}}) \tag{H.4}$$

where the coefficient $1-\alpha_j$ applies to q_t, coefficients $\frac{\alpha_j}{K_j-1}$ apply to the firm's own competitors and the 0 coefficients apply to firms in other industries. Plugging these into Equation (H.2) and setting $\beta = 0$ we obtain the following Problem for the firm at time t:

$$\max_{S_t \in \mathbb{S}^t, p_t: S^t \to \mathbb{R}} \mathbb{E}[-\frac{B_j}{2}(p_t - \mathbf{w}'\vec{x}_t)^2 - \omega \mathcal{I}(X^t(\varsigma); S^t | S^{t-1}) | S^{-1}] \tag{H.5}$$

First, we can see that for any choice of $S_t \in \mathbb{S}^t$, the optimal pricing strategy is given by:

$$p_t(S^t) = \mathbf{w}'\mathbb{E}[\vec{x}_t | S^t] \tag{H.6}$$

$$\implies -\mathbb{E}[\frac{B_j}{2}(p_t - \mathbf{w}'\vec{x}_t(\varsigma)) | S^{-1}] = -\frac{B_j}{2}\mathbb{E}[\mathbb{V}\text{ar}(\mathbf{w}'\vec{x}_t(\varsigma) | S^t) | S^{-1}] \tag{H.7}$$

Moreover, by the chain rule of mutual information, we have the following decomposition of the firm's cost of information:

$$\omega \mathcal{I}(X^t(\varsigma); S^t | S^{t-1}) = \omega \mathcal{I}(X^t(\varsigma); S_t, S^{t-1} | S^{t-1}) = \omega \mathcal{I}(X^t(\varsigma); S_t | S^{t-1}) \tag{H.8}$$

$$= \omega \mathcal{I}(X^{t-1}(\varsigma), \vec{x}_t(\varsigma); S_t | S^{t-1}) \tag{H.9}$$

$$= \omega \mathcal{I}(\vec{x}_t(\varsigma); S_t | S^{t-1}) + \omega \mathcal{I}(X^{t-1}(\varsigma); S_t | S^{t-1}, \vec{x}_t(\varsigma)) \tag{H.10}$$

Thus, the firm's problem with $\beta = 0$ at time t can be written as:

$$\max_{S_t \in \mathbb{S}^t} -\mathbb{E}[\underbrace{\frac{B_j}{2}\mathbf{w}'\mathbb{V}\text{ar}(\vec{x}_t(\varsigma) | S^t)\mathbf{w}}_{\text{losses from mispricing}} + \underbrace{\omega \mathcal{I}(\vec{x}_t(\varsigma); S_t | S^{t-1})}_{\text{cost of info. about } \vec{x}_t \text{ conditional on } S^{t-1}} + \underbrace{\omega \mathcal{I}(X^{t-1}(\varsigma); S_t | S^{t-1}, \vec{x}_t(\varsigma))}_{\text{cost of info. about } X^{t-1} \text{ conditional on } S^{t-1}, \vec{x}_t} | S^{-1}] \tag{H.11}$$

Our first observation about this problem is that it is optimal to choose S_t such that the third term (cost of info. about X^{t-1} conditional on $S^{t-1}, \vec{x}_t$) is zero; i.e., choose a signal that is not informative about past fundamentals and prices *conditional* on today's prices and fundamentals. To see why, suppose that this term is strictly positive so that S_t contains some information about X^{t-1} that is *independent* of $\vec{x}_t, S^{t-1}$. But that cannot be optimal because one can construct a new signal that has the same amount of information about $\vec{x}_t(\varsigma)$ conditional on S^{t-1} but less information about $X^{t-1}(\varsigma)$ conditional on $S^{t-1}, \vec{x}_t(\varsigma)$. Such a signal would imply the same losses from mispricing but would economize on irrelevant information about $X^{t-1}(\varsigma)$ that are not relevant for predicting $\vec{x}_t(\varsigma)$. Thus, the firm's problem reduces to:

$$\max_{S_t \in \mathbb{S}^t} -\mathbb{E}[\frac{B_j}{2}\mathbf{w}'\mathbb{V}\text{ar}(\vec{x}_t(\varsigma) | S^t)\mathbf{w} + \omega \mathcal{I}(\vec{x}_t(\varsigma); S_t | S^{t-1}) | S^{-1}] \tag{H.12}$$

Now, given S^{t-1} and $\vec{x}_t(\varsigma)$, let $\Sigma_{t|t-1} \equiv \mathbb{V}\text{ar}(\vec{x}_t(\varsigma) | S^{t-1})$. Assume, without loss of generality, that $\Sigma_{t|t-1}$

is invertible.[54] Moreover, notice that by restricting the strategies to be among Gaussian signals as well as the fact that q_t itself is a Gaussian process, for any non-zero signal $S_t \in \mathbb{S}^t$, we have:

$$\omega \mathcal{I}(S_t, \vec{x}_t(\varsigma)|S^{t-1}) = \frac{\omega}{2}\ln(1 - \mathbf{z}_t' \Sigma_{t|t-1}^{-1} \mathbf{z}_t),$$

where $\mathbf{z}_t \equiv \frac{\mathbb{C}\text{ov}(S_t, \vec{x}_t(\varsigma)|S^{t-1})}{\sqrt{\mathbb{V}\text{ar}(S_t|S^{t-1})}}$. Moreover, notice that firm's losses from mispricing become

$$\mathbb{V}\text{ar}(\mathbf{w}'\vec{x}_t(\varsigma)|S_{j,k}^{t-1}, S_{j,k,t}) = \mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w} - (\mathbf{w}'\mathbf{z}_t)^2.$$

Since both the cost of information and losses from mispricing are functions of the signal through $\mathbf{z}_t$, the firm can directly choose $\mathbf{z}_t$ (as long as there is a signal in $\mathbb{S}^t$ that induces that covariance vector $\mathbf{z}_t$, which corresponds to the no-forgetting constraint). Assuming that the no-forgetting constraint does not bind for the moment, the first order condition for $\mathbf{z}_t$ is:

$$B_j(\mathbf{w}'\mathbf{z}_t^*)\mathbf{w} = \omega \frac{\Sigma_{t|t-1}^{-1}\mathbf{z}_t^*}{1 - \mathbf{z}_t^{*\prime}\Sigma_{t|t-1}^{-1}\mathbf{z}_t^*} \tag{H.13}$$

multiplying this FOC with $\mathbf{z}_t^*$ and $\mathbf{w}$ from left, implies that

$$(\mathbf{w}'\mathbf{z}_t^*)^2 = \frac{\omega}{B_j}\frac{\mathbf{z}_t^{*\prime}\Sigma_{t|t-1}^{-1}\mathbf{z}_t^*}{1 - \mathbf{z}_t^{*\prime}\Sigma_{t|t-1}^{-1}\mathbf{z}_t^*}, \qquad \mathbb{V}\text{ar}(p^*(\vec{x}_t(\varsigma))|S^{t-1}) = \mathbf{w}'\Sigma_{t|t-1}\mathbf{w} = \frac{\omega}{B_j}\frac{1}{1 - \mathbf{z}_t^{*\prime}\Sigma_{t|t-1}^{-1}\mathbf{z}_t^*} \tag{H.14}$$

Combining these two equations, we have:

$$(\mathbf{w}'\mathbf{z}_t^*)^2 = \mathbf{w}'\Sigma_{t|t-1}\mathbf{w} - \frac{\omega}{B_j} \tag{H.15}$$

Since the left-hand side is a positive number, this requires that

$$\mathbb{V}\text{ar}(p^*(\vec{x}_t(\varsigma))|S^{t-1}) = \mathbf{w}'\Sigma_{t|t-1}\mathbf{w} \geq \frac{\omega}{B_j} \tag{H.16}$$

which is what is required for the no-forgetting constraint to not bind; i.e., the prior uncertainty of the firm about its ideal price, $\mathbf{w}'\Sigma_{t|t-1}\mathbf{w}$, needs to be large enough so that it pays attention to it. Otherwise, the optimal signal has zero covariance with $\vec{x}_t(\varsigma)$. Therefore, the optimal covariance is proportional to $\Sigma_{t|t-1}\mathbf{w}$:

$$\mathbf{z}_t^* = \frac{\max\{\sqrt{\mathbf{w}'\Sigma_{t|t-1}\mathbf{w} - \frac{\omega}{B_j}}, 0\}}{\mathbf{w}'\Sigma_{t|t-1}\mathbf{w}} \times \Sigma_{t|t-1}\mathbf{w} \tag{H.17}$$

where the first term is a scalar that depends on the cost and benefit parameters ω, B_j. The last step is to characterize a signal $S_t \in \mathbb{S}^t$ that implies this optimal covariance. To see this, let

$$S_t^* \equiv \mathbf{w}'\vec{x}_t(\varsigma) + e_t = (1-\alpha)q_t + \alpha\frac{1}{K-1}\sum_{l \neq k} p_{j,l,t}(\varsigma) + e_t.$$

where e_t is a Gaussian noise independent of $\vec{x}_t(\varsigma)$. It is straight forward to show that this signal implies $\mathbf{z}_t^*$ for an appropriately chosen variance for e_t.[55]

[54] To see why this is without loss of generality, note that if $\Sigma_{t|t-1}$ is not invertible, then there are elements in $\vec{x}_t(\varsigma)$ that are colinear conditional on $S_{j,k}^{t-1}$, in which case knowing about one completely reveal the other; this means we can reduce $\vec{x}_t(\varsigma)$ to its orthogonal elements without limiting the signal choice of the agent.

[55] Here one needs to define $\mathbb{V}\text{ar}(e_t) = \infty$ to correspond to the case where $\mathbf{z}_t^* = 0$, which is a well-defined limit.

Proof of Proposition 6

The independence of strategic complementarity α_j from j follows from the symmetry in the number of competitors across industries. Moreover, in the stationary equilibrium capacity is time-invariant because it only depends on the underlying parameters and the variances of subjective beliefs, which are constant under the steady-state Kalman filter. Symmetric equilibrium also implies that optimal capacities are also symmetric across all firms; so $\kappa_{j,k,t}=\kappa\geq 0$. To see that $\kappa>0$, suppose that in the equilibrium $\kappa=0$. Then firms are not acquiring any information about the prices of their competitors and the monetary policy shocks. But monetary policy shocks have a unit root which under the assumption that $\kappa=0$ implies that firms' uncertainty about their optimal price, which is proportional to their losses from imperfect information, is growing linearly over time and exceeds any finite upper-bound. Now, consider an information acquisitoin strategy that sets $\kappa=\epsilon>0$. It follows that firms' losses under this strategy is bounded above by $\mathcal{O}(\frac{1}{\epsilon})$ which dominates $\kappa=0$. Thus, in the stationary equilibrium, $\kappa>0$.

Now, from the proof of Proposition 5, recall that in the equilibrium, for all $(j,k)\in J\times K$, $p_{j,k,t}(S^t_{j,k})=\mathbf{w}'\mathbb{E}[\vec{x}_{j,k,t}(\varsigma)|S^t_{j,k}]$ where $S^t_{j,k}=(S^{t-1}_{j,k},S_{j,k,t})$ and

$$S_{j,k,t}=(1-\alpha)q_t+\alpha\frac{1}{K-1}\sum_{l\neq k}p_{j,l,t}(S^t_{j,l})+e_{j,k,t}$$

From Kalman filtering

$$\begin{aligned}\mathbf{w}'\mathbb{E}[\vec{x}_{j,k,t}(\varsigma)|S^t_{j,k}] &= \mathbb{E}[\mathbf{w}'\vec{x}_{j,k,t}(\varsigma)|S^{t-1}_{j,k}] \\ &+ \frac{\mathbf{w}'\mathbb{C}\mathrm{ov}(S_{j,k,t},\vec{x}_t(\varsigma))}{\mathbb{V}\mathrm{ar}(S_{j,k,t}|S^{t-1}_{j,k})}(S_{j,k,t}-\mathbb{E}[S_{j,k,t}|S^{t-1}_{j,k}]).\end{aligned}$$

Notice from the proof of Proposition 5 that $\frac{\mathbf{w}'\mathbb{C}\mathrm{ov}(S_{j,k,t},\vec{x}_{j,k,t}(\varsigma))}{\mathbb{V}\mathrm{ar}(S_{j,k,t}|S^{t-1}_{j,k})}=\frac{\lambda}{\mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w}}\mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w}=\lambda$. Thus, using $p_{j,k,t}$ as shorthand for $p_{j,k,t}(S^t_{j,k})$, $p_{j,k,t}=(1-\lambda)\mathbb{E}[S_{j,k,t}|S^{t-1}_{j,k}]+\lambda S_{j,k,t}$. Finally, notice that $p_{j,k,t-1}=\mathbb{E}[S_{j,k,t-1}|S^{t-1}_{j,k}]$. Subtract this from both sides of the above equation to get $\pi_{j,k,t}\equiv p_{j,k,t}-p_{j,k,t-1}=(1-\lambda)\mathbb{E}[\Delta S_{j,k,t}|S^{t-1}_{j,k}]+\lambda(S_{j,k,t}-p_{j,k,t-1})$, where $\Delta S_{j,k,t}=S_{j,k,t}-S_{j,k,t-1}$. Subtract $\lambda\pi_{j,k,t}$ from both sides and divide by $(1-\lambda)$ to get $\pi_{j,k,t}=\mathbb{E}[\Delta S_{j,k,t}|S^{t-1}_{j,k}]+\frac{\lambda}{1-\lambda}(S_{j,k,t}-p_{j,k,t})$. Averaging, this equation over all firms gives us the Phillips curve:

$$\overline{\mathbb{E}^{j,k}_{t-1}[\Delta S_{j,k,t}]} \equiv \frac{1}{JK}\sum_{(j,k)\in J\times K}\mathbb{E}[\Delta S_{j,k,t}|S^{t-1}_{j,k}]=(1-\alpha)\overline{\mathbb{E}^{j,k}_{t-1}[\Delta q_t]}+\alpha\overline{\mathbb{E}^{j,k}_{t-1}[\pi_{j,-k,t}]}.$$

where $\pi_{j,-k,t}\equiv\frac{1}{K-1}\sum_{l\neq k}(p_{j,l,t}-p_{j,l,t-1})$ is the average price change of all others in industry j except k. Moreover,

$$\frac{1}{JK}\sum_{(j,k)\in J\times K}(S_{j,k,t}-p_{j,k,t}) = (1-\alpha)q_t+\underbrace{\frac{\alpha}{JK}\sum_{(j,k)\in J\times K}\frac{1}{K-1}\sum_{l\neq k}p_{j,l,t}-\frac{1}{JK}\sum_{(j,k)\in J\times K}p_{j,k,t}}_{=\frac{\alpha-1}{JK}\sum_{(j,k)\in J\times K}p_{j,k,t}}.$$

The last term asymptotically converges to zero as $J\to\infty$ as mistakes are orthogonal across sectors— $e_{j,k,t}\perp p_{m,l,t},\forall m\neq j$. Now, define $p_t\equiv\frac{1}{JK}\sum_{(j,k)\in J\times K}p_{j,k,t}$, and recall that $q_t=p_t+y_t$. Therefore,

$\frac{1}{JK}\sum_{(j,k)\in J\times K}(S_{j,k,t}-p_{j,k,t})=(1-\alpha)y_t$. Finally, define aggregate inflation as the average price change in the economy, $\pi_t \equiv \frac{1}{JK}\sum_{(j,k)\in J\times K}\pi_{j,k,t}$. Plugging these into the expression above we get

$$\pi_t=(1-\alpha)\overline{\mathbb{E}_{t-1}^{j,k}[\Delta q_t]}+\alpha\overline{\mathbb{E}_{t-1}^{j,k}[\pi_{j,-k,t}]}+(1-\alpha)\frac{\lambda}{1-\lambda}y_t.$$

Finally, notice that $\frac{\lambda}{1-\lambda}=\frac{1-e^{-2\kappa}}{e^{-2\kappa}}=e^{2\kappa}-1$.

I Calibration Details

I.1. Calibration of the Benchmark Model

This section discusses the calibration of several model parameters in detail.

Elasticity of substitution. A usual approach in monopolistic competition models is to choose η to match an average markup given by $\frac{\eta}{\eta-1}$. In the oligopolistic competition model, markups depend on the number of competitors and in the steady-state are given by

$$\mu_j=1+\frac{1}{(\eta-1)(1-K_j^{-1})} \tag{I.1}$$

where K_j is the number of competitors in j. The survey elicits firms' markups by asking the following question: *"Considering your main product line or main line of services in the domestic market, by what margin does your sales price exceed your operating costs (i.e., the cost material inputs plus wage costs but not overheads and depreciation)? Please report your current margin as well as the historical or average margin for the firm."* The average markup reported by firms in the sample is 1.3 and varies from 1.1 to 1.6. These values are in the plausible range of markups measured in the literature for the US. Given this measure of markups, I run the analogous regression to Equation (I.1) and set $\eta=12$ to match the coefficient on $\frac{1}{1-K_j^{-1}}$ in Column (2) of Table I.1, which reports the result of this regression. This value is well in line with the values used in the literature for the US.

Table I.1: Calibration of η

	(1) Average Markup		(2) Average Markup	
$1/(1-K^{-1})$	0.107	(0.016)	0.089	(0.018)
Firm age			0.000	(0.000)
Manufacturing			0.037	(0.007)
Professional and Financial Services			0.166	(0.007)
Trade			0.027	(0.007)
Other			-0.031	(0.044)
Constant	1.205	(0.018)	1.140	(0.021)
Observations	3152		3152	

Notes: Column (1) of the table reports the result of regressing the average markups of firms on $1/(1-K_j^{-1})$ in the first wave of the survey from Coibion et al. (2018). Column (2) controls for industry fixed effects shown in the table as well as firm age. The coefficient on $1/(1-K_j^{-1})$ corresponds to $1/(\eta-1)$ in the model.

Curvature of the production function. Given the empirical distribution of the number of firms, $\mathcal{K}$, and the elasticity of substitution, $\eta = 12$, I set $\gamma = 0.945$ ($1/(1+\gamma) = 0.514$) to match the average degree of strategic complementarity $\bar{\alpha} = 0.817$ from Table A.1. Given this value, the elasticity of output to labor in the model is 0.514. This is consistent with calibrations of this parameter for the U.S. if we were to calibrate it to the labor share of income in the U.S. data (see e.g. Bilal, Engbom, Mongey, & Violante, 2019, where the targeted value for the U.S. is 0.518).[56]

Persistence and variance of shocks to nominal demand. I calibrate $\rho = 0.707$ to match the persistence of the growth of nominal GDP in New Zealand for post-1991 and pre COVID-19 data.[57] Nonetheless, the model is not very sensitive to this parameter in this range and I present results for an alternative value of $\rho = 0.23$ in Section 5.4.

Given the quarterly persistence, I then set $\sigma_u = 0.011$ to match the unconditional standard deviation of quarterly nominal GDP growth.[58] Nonetheless, since monetary policy shocks are the only shocks in the model, the standard deviation of all variables – including endogenous non-fundamental shocks – are scaled by the standard deviation of the innovations to nominal demand. Accordingly, in my counterfactual comparisons, I will mainly focus on numbers relative to a benchmark so that the reported relative numbers are independent of this scale.[59]

I.2. Calibration of the Monopolistic Competition Model

In Section 5, I compare the calibrated benchmark model to a monopolistic competition model with the same average degree of strategic complementarity. In this section, I discuss the calibration of this monopolistic competition model. Since changing K affects both the degree of strategic complementarity, α_j, as well as the curvature of the profit function, B_j in Equation (28), by taking the limit $K \to \infty$ in the benchmark model, we would inevitably alter both the curvature and the degree of strategic complementarity. To avoid this and create the degree of freedom that allows for keeping the strategic complementarity fixed as $K \to \infty$, in the monopolistic competition model I replace the within-industry CES aggregators with a Kimball aggregator. In particular, I use the oligopolistic Kimball aggregator derived in Appendix G which keeps the demand elasticities the same at $\varepsilon_j^D = \eta - (\eta - 1)K^{-1}$ but introduces the parameter ζ that controls the degree of strategic complementarity. The degree of strategic complementarity in this case

[56]Although we have not explicitly modeled capital, one could think of the production function of firms as one with constant returns to scale in capital and labor, where capital is exogenously fixed.

[57]This coefficient is obtained by regressing the annual log-growth of nominal GDP in New Zealand on one lag where I obtain a yearly persistence of 0.25. I then convert this to the quarterly persistence through $\rho = 0.25^{1/4} = 0.707$. I restrict the time series to post 1991 data to be consistent with New Zealand's shift in monetary policy towards inflation targeting in that time frame. I also restrict the data to pre 2020 to exclude the COVID-19 period from the sample.

[58]The unconditional standard deviation is given by $\frac{\sigma_u}{\sqrt{1-\rho^2}}$ in the model which is 0.0154 in the data.

[59]This is due to potential concerns in matching the unconditional volatility. Calibrating the standard deviation needs to be done on the part of nominal demand that is driven by monetary policy shocks. In the US one can calibrate this variance by projecting nominal demand on known monetary policy shock series, such as Romer and Romer (2004) shocks, and fitting an AR(1) to the predicted series (See, for instance, Midrigan (2011)). For the case of New Zealand, however, this becomes a complication since, as far as I know, there is no unanimously agreed-upon series for monetary shocks.

and when $K \to \infty$, is then given by

$$\lim_{K\to\infty} \alpha_K = \zeta \frac{1}{1+\gamma\eta} + (1-\zeta)\frac{\gamma(\eta-1)}{1+\gamma\eta} \tag{I.2}$$

The new parameter ζ allows us to match the degree of strategic complementarity to any target in the monopolistic competition model. Importantly, ζ does not directly affect anything else in the model and only shows up in the expression for α.

J Symmetric Stationary Equilibria and Solution Method

This appendix has four subsections. Appendix J.1 defines the notion of a symmetric stationary equilibrium for the dynamic model and discusses some of its properties. Appendix J.2 derives the equations that need to hold in such an equilibrium. It further characterizes this equilibrium as the fixed point of a mapping on a set of lag polynomials for the evolution of prices, given a Markov state space approximation that is used for solving the dynamic rational inattention problem of firms. Appendix J.3 outlines the main algorithm that I use for finding this fixed point using an "integrated moving average (MA)" state space approximation (Algorithm 1). It concludes with a description of how this method is implemented in the replication package of the paper, which is publicly available at https://doi.org/10.7910/DVN/AO6C85. Finally, Appendix J.4 examines the robustness of Algorithm 1 by replacing the integrated MA approximation with an alternative one based on an ARMA approximation (Algorithm 2), as in Maćkowiak, Matějka, and Wiederholt (2018), and shows that the results delivered by these two alternative algorithms are numerically identical by comparing the maximum distance between their implied IRFs as well as replicating the main quantitative results in Tables 4 to 6.

J.1. Definition of Symmetric Stationary Equilibria

To define the notion of an equilibrium for the dynamic game, let $(S_{j,k}^{-1})_{(j,k)\in(J\times K)}$ denote the initial information sets that firms take as given at time 0. Moreover, let $\varsigma_{j,k}$ denote a strategy profile for any firm j,k in the economy, which consists of choices of signals over time along with pricing strategies for every period that maps the firm's information set at that time to a price:

$$\varsigma_{j,k} = (S_{j,k,t} \subset \mathbb{S}^t, S_{j,k}^t = S_{j,k}^{t-1} \cup S_{j,k,t}, p_{j,k,t} : S_{j,k}^t \to \mathbb{R})_{t\geq 0} \qquad S_{j,k}^{-1} \text{ given.} \tag{J.1}$$

We start by revisiting the problem of a single firm given a set of strategy profiles for other firms. Fix a firm j,k, and (with a slight abuse of notation) let $\varsigma \equiv (\varsigma_{l,m})_{(l,m)\neq(j,k)}$ a set of strategies (not necessarily equilibrium strategies) for all other firms. Define $\vec{x}_{j,k,t}(\varsigma) \equiv (q_t, p_{l,m,t}(S_{l,m}^t))_{(l,m)\neq(j,k)}$ and $X_{j,k}^t(\varsigma) \equiv \{\vec{x}_{j,k,\tau}(\varsigma) : 0 \leq \tau \leq t\}$. Note that under strategy profile ς, $X_{j,k}^t(\varsigma)$ has a stochastic (but potentially time-varying) process that is exogenous to firm j,k and is taken as given by that firm—this process can be time-varying because ς can be such that other firms are changing their pricing strategies over time beyond what is implied by shocks alone.

In proof of Proposition 6, in Equation (H.2), we showed that with this notation, given a ς, firm j,k's problem under an initial information set $S_{j,k}^{-1}$ can be cast in the following form:

$$\max_{\{S_{j,k,t}\subset \mathbb{S}_{j,k}^t, p_{j,k,t}(S^t)\to\mathbb{R}\}_{t\geq 0}} -\sum_{t=0}^{\infty}\beta^t \mathbb{E}[\frac{1}{2}B_j(p_{j,k,t}-p^*(\vec{x}_{j,k,t}(\varsigma)))^2+\omega\mathcal{I}(X_{j,k}^t(\varsigma);S_{j,k}^t|S_{j,k}^{t-1})|S_{j,k}^{-1}] \tag{J.2}$$

$$\text{subject to}\quad S_{j,k}^t = S_{j,k}^{t-1}\cup S_{j,k,t},\quad \forall t\geq 0,\quad S_{j,k}^{-1}\text{ given},\quad p^*(x_{j,k,t}(\varsigma))\equiv(1-\alpha_j)q_t+\alpha_j p_{j,-k,t}(\varsigma)$$

With this specification of a firm's problem at hand, the following definition extends the notion of a pure strategy Gaussian equilibrium that we defined for the static game in Definition 1 to the dynamic case:

Definition 2. A pure strategy Gaussian equilibrium is a collection of initial information sets, $(S_{j,k}^{-1})_{(j,k)\in J\times K}$ along with a collection of strategies for firms $(\varsigma_{j,k})_{(j,k)\in(J\times K)}$ such that (1) given these strategies and initial information sets, no firm j,k has the incentive to deviate from $\varsigma_{j,k}$ according to the objective defined in Equation (J.2), and (2) $(q_t, p_{j,k,t}(S_{j,k}^t))_{t\geq 0}^{(j,k)\in(J\times K)}$ is a multivariate Gaussian process. Moreover, we call such a pair of initial information sets and strategies, $(S_{j,k}^{-1},\varsigma_{j,k})_{(j,k)\in(J\times K)}$, a *symmetric stationary Gaussian equilibrium*, if they also satisfy the following additional conditions:

1. **Symmetry:** the pricing strategies of firms within all sectors with $K_j = K\in \text{Supp}(\mathcal{K})$ competitors are independent of firms' identity (index) and only depend on their information sets:
$$\forall t\geq 0, \forall S^t\subset \mathbb{S}^t, \forall (j,k),(l,m)\in(J\times K), K_j = K_l:\qquad p_{j,k,t}(S^t) = p_{l,m,t}(S^t) \tag{J.3}$$
2. **Stationarity:** the pricing strategies of all firms depend on time only through their history of signals and not on the time index itself:
$$\forall t,h\geq 0, \forall S^t\in\mathbb{S}^t, \forall (j,k)\in(J\times K):\qquad p_{j,k,t}(S^t) = p_{j,k,h}(S^t) \tag{J.4}$$

Discussion. To clarify the restrictions of symmetry and stationarity, let us note that since, in a Gaussian equilibrium, the ideal price of firm j,k, $p^*(x_{j,k,t}(\varsigma))$ in Equation (J.2) is also Gaussian (because it is the sum of Gaussian processes), firm j,k's optimal price, $p_{j,k,t}(S_{j,k}^t) = \mathbb{E}(p^*(\vec{x}_{j,k,t})|S_{j,k^t})$, will be a linear function of the history of its signals:

$$p_{j,k,t}(S_{j,k}^t) = \sum_{\tau\geq 0}\delta_{j,k,t}^{\tau} S_{j,k,t-\tau} \tag{J.5}$$

where coefficients $(\delta_{j,k,t})_{\tau\geq 0}$ are determined by its optimal Kalman filtering problem. In this context, symmetry requires that for any two firms (j,k) and (l,m) in oligopolies with $K_j = K_l = K\in \text{Supp}(\mathcal{K})$ competitors

$$\delta_{j,k,t}^{\tau} = \delta_{l,m,t}^{\tau} = \delta_{K,t}^{\tau},\quad \forall t\geq 0, \forall\tau\geq 0 \tag{J.6}$$

Furthermore, stationarity requires that

$$\delta_{K,t}^{\tau} = \delta_{K,h}^{\tau} = \delta_{K}^{\tau},\quad \forall t,h\geq 0, \forall\tau\geq 0 \tag{J.7}$$

An immediate observation is that stationarity can arise only for specific initial information sets. For instance, if $\delta_{j,k,t}^{\tau}\neq 0, \forall\tau\geq 0$ which is usually the case in imperfect information models with endogenous signals (usually referred to as the "infinite regress property"), a necessary condition for the equilibrium to be stationary is that there should be countably infinitely many signals in $S_{j,k}^{-1}$ for all $(j,k)\in(J\times K)$.

Moreover, these infinitely many signals should be such that given this history and others' strategies, a firm j,k will continue to choose signals over time that induce similar optimal filtering behavior. This is a well-known property of the equilibrium in rational inattention pricing models with monopolistic competition, pointed out by Maćkowiak and Wiederholt (2009). In such models with monopolistic competition, the game specified here is a mean-field game with infinitely many firms, where every infinitesimal firm takes a *stationary* process for the average price of all firms as given, and chooses its own pricing plan to maximize its profits. It is then assumed that each firm receives infinitely many signals at the initial period so that its own optimal pricing strategy is also stationary. The equilibrium in that setting is then a *fixed point* where the stationary process of the average price is consistent with the optimal pricing plan of each individual firm.

In this sense, the definition above formalizes and extends this fixed point notion to the case of the game with many sectors with finite and potentially different numbers of competitors: each firm j,k takes as given that its competitors are following a stationary pricing strategy, and chooses its own pricing strategy to maximize its profits. We then require that this firm starts from an initial information set that is such that its optimal pricing strategy is stationary (as in Maćkowiak and Wiederholt (2009), this is similar to assuming that once the firm has solved its problem at time 0, it will receive infinitely many signals such that its initial information set induces a stationary filtering behavior going forward).

The equilibrium is then again a fixed point for these strategies and initial information sets such that the stationary strategy of other firms is consistent with the optimal pricing strategy of each individual firm, given its own initial information set.[60]

Finally, it is worth noting that while a symmetric stationary equilibrium implies that firms' pricing strategies remain time-invariant over time (and in this sense constitutes a "steady state"), it does not necessarily need to emerge as the limiting steady state of the rational inattention game starting from *any* arbitrary set of initial information sets. But *conditional* on a steady state emerging from a set of initial information sets for the game, such a steady state must be a stationary equilibrium. In this sense, the definition of stationarity above is a necessary condition for the existence of a steady state for the game.[61]

J.2. Solution Method and Characterization of the Symmetric Stationary Equilibrium

Suppose the pair $(S_{j,k}^{-1}, \varsigma_{j,k})_{(j,k)\in(J\times K)}$ is a Gaussian symmetric stationary equilibrium. This section outlines a two-stage procedure for finding the fixed point described above. In the first stage, I solve for a firm's optimal pricing strategy, given that others play a stationary strategy. In the second stage, I derive the conditions for the symmetric stationary equilibrium fixed point. I use these conditions in the next

[60]For a broader context on this definition, it might also be useful to consider the following analogy with heterogeneous agent models such as Bewley-Huggett-Aiyagari models. In these models, policy functions are typically time-dependent given the initial distribution of endogenous variables, but a stationary equilibrium can be defined as the one that arises under a particular "stationary initial distribution" that induces optimal policy functions that are time-independent.

[61]This is again analogous to the case of heterogeneous agent models where a stationary equilibrium is a necessary condition for the existence of a steady state but such a steady state might or might not emerge as the limiting steady state of such an economy starting from *any* initial distribution.

two subsections to outline two iteration algorithms for how one can update the guess for the symmetric stationary strategies of other firms in the equilibrium and repeat the process until convergence.

Stage 1: Solving a single firm's problem. Fix a firm $(j,k) \in (J \times K)$ and suppose that each of its competitors price according to a symmetric stationary pricing strategy so that their price can be decomposed to its projection on fundamental shocks and orthogonal residuals:

$$p_{j,l,t} = \psi_{j,u}(L)u_t + \psi_{j,v}(L)v_{j,l,t}, \quad \forall l \neq k \tag{J.8}$$

where $\psi_{j,u}(L)$ and $\psi_{j,v}(L)$ are lag polynomials and u_t and $v_{j,l,t}$ are Gaussian innovations to money growth and the mistake of firm j,l (due to the rational inattention errors in its signals), respectively. It then follows that the average price of firm j,k's competitors, $p_{j,-k,t}$, follows:

$$p_{j,-k,t} = \frac{1}{K_j - 1}\sum_{l \neq k} p_{j,l,t} = \psi_{j,u}(L)u_t + \psi_{j,v}(L)v_{j,-k,t}, \quad v_{j,-k,t} = \frac{1}{K_j - 1}\sum_{l \neq k} v_{j,l,t} \tag{J.9}$$

where we normalize the scales of the lag polynomials such that $\mathbb{V}\text{ar}(u_t) = \mathbb{V}\text{ar}(v_{j,-k,t}) = \sigma_u^2$. It then follows that the firm j,k's ideal price is given by:

$$p^*_{j,k,t} = (1-\alpha_j)q_t + \alpha_j p_{j,-k,t} \tag{J.10}$$

$$= ((1-\alpha_j)\psi_{q,u}(L) + \alpha_j \psi_{j,u}(L))u_t + \alpha_j \psi_{j,v}(L)v_{j,-k,t} \tag{J.11}$$

where $\psi_{q,u}(L)$ is the lag polynomial that maps innovations to money growth to nominal demand:

$$\Delta q_t = \rho \Delta q_{t-1} + u_t \Longrightarrow q_t = \frac{u_t}{(1-\rho L)(1-L)} \Longrightarrow \psi_{q,u}(L) = \frac{1}{(1-\rho L)(1-L)} \tag{J.12}$$

Now, for ease of notation, let us define:

$$x^u_{j,k,t} \equiv ((1-\alpha_j)\psi_{q,u}(L) + \alpha_j \psi_{j,u}(L))u_t \tag{J.13}$$

$$x^v_{j,k,t} \equiv \alpha_j \psi_{j,v}(L)v_{j,-k,t} \tag{J.14}$$

so that we can write

$$p^*_{j,k,t} = x^u_{j,k,t} + x^v_{j,k,t} \tag{J.15}$$

where $x^u_{j,k,t}$ is the projection of the ideal price on the history of monetary shocks. $x^v_{j,k,t}$ is then the residual, representing the variation in the firm's ideal price that is induced by its competitors' mistakes in their signals, which are potentially correlated but are independent of monetary shocks.[62]

In the next step, to map this problem to the Gaussian dynamic rational inattention problem as in Maćkowiak et al. (2018) or Afrouzi and Yang (2019), we approximate the processes $\psi_{j,u}(L)u_t$ and $\psi_{j,v}(L)v_{j,-k,t}$ such that they can be written as a linear function of a multivariate process with a Markov state space representation.[63] In particular, suppose there exist Markov Gaussian processes $\xi^u_t \in \mathbb{R}^n$ and $\xi^v_t \in \mathbb{R}^m$ for some $n,m \in \mathbb{N}$ such that

$$x^u_{j,k,t} \approx H'_{j,u}\xi^u_{j,t}, \quad \xi^u_{j,t} = A_{j,u}\xi^u_{j,t-1} + Q_{j,u}u_t, \quad A_{j,u} \in \mathbb{R}^{n\times n}, \quad Q_{j,u} \in \mathbb{R}^n \tag{J.16}$$

[62] In monopolistic competition models with rational inattention, it is assumed that $x^v_{j,k,t} = 0$, but that is not the case here because of a finite number of firms in every sector.

[63] See, e.g., Han, Tan, and Wu (2022) for a proof that such processes can be approximated in this way with arbitrary accuracy.

$$x^v_{j,k,t} \approx H'_{j,v}\xi^v_{j,k,t}, \quad \xi^v_{j,t} = A_{j,v}\xi^v_{j,t-1} + Q_{j,v}v_{j,-k,t}, \quad A_{j,v} \in \mathbb{R}^{m\times m}, \quad Q_{j,v} \in \mathbb{R}^m \tag{J.17}$$

Below, I will discuss two different approaches to perform this approximation and obtain the matrices $H_{j,u}, H_{j,v}, A_{j,u}, A_{j,v}, Q_{j,u}, Q_{j,v}$, one using an integrated MA truncation and another using an ARMA approximation as in Maćkowiak et al. (2018) and show both methods deliver numerically identical results. For this section, however, let us take the state space representations of $\xi^u_{j,t}$ and $\xi^v_{j,k,t}$ as given. We can define the augmented state $\xi_{j,k,t} = \begin{bmatrix} \xi^u_{j,t} \\ \xi^v_{j,k,t} \end{bmatrix}$ which has the following stationary Markov state space representation:

$$\xi_{j,k,t} = \underbrace{\begin{bmatrix} A_{j,u} & 0 \\ 0 & A_{j,v} \end{bmatrix}}_{A_j \in \mathbb{R}^{(n+m)\times(n+m)}} \xi_{j,k,t-1} + \underbrace{\begin{bmatrix} Q_{j,u} & 0 \\ 0 & Q_{j,v} \end{bmatrix}}_{Q_j \in \mathbb{R}^{(n+m)\times 2}} \underbrace{\begin{bmatrix} u_t \\ v_{j,-k,t} \end{bmatrix}}_{\equiv \epsilon_{j,k,t} \sim N(0,\sigma_u^2 I_2)} \tag{J.18}$$

Therefore, $p^*_{j,k,t}$ is approximated by $H'_j\xi_{j,k,t}$ where $H_j = \begin{bmatrix} H_{j,u} \\ H_{j,v} \end{bmatrix}$.

With this state space representation, the problem in Equation (J.2) is analogous to the one derived in Lemma 2.4 (Proposition 1 in October 2019 version) in Afrouzi and Yang (2019). Following the same steps in the proof of that problem, we can then write the firm's problem as

$$\max_{\{\Sigma_{j,t|t} \succsim 0\}_{t\geq 0}} -\frac{1}{2}\sum_{t=0}^{\infty} \left[B_j tr(H_j H'_j \Sigma_{j,t|t}) + \omega \ln\left(\frac{|\Sigma_{j,t|t-1}|}{|\Sigma_{j,t|t}|} \right) \right] \tag{J.19}$$

$$s.t. \quad \Sigma_{j,t+1|t} = A_j \Sigma_{j,t|t} A'_j + \sigma_u^2 Q_j Q'_j, \quad \forall t \geq 0 \tag{J.20}$$

$$\Sigma_{j,t|t-1} - \Sigma_{j,t|t} \succsim 0 \quad \forall t \geq 0, \qquad \Sigma_{0|-1} \equiv \mathbb{V}\text{ar}(\xi_{j,k,0}|S^{-1}_{j,k}) \tag{J.21}$$

where $\succsim 0$ denotes positive semi-definiteness. Here $\Sigma_{j,t|t-1} = \mathbb{V}\text{ar}(\xi_{j,k,t}|S^{t-1}_{j,k})$ and $\Sigma_{t|t} = \mathbb{V}\text{ar}(\xi_{j,k,t}|S^t_{j,k})$ denote the prior and posterior covariance matrices of the firms' beliefs about $\xi_{j,k,t}$ at time t given their information sets at time $t-1$ and t, respectively. Moreover, $\Sigma_{j,0|-1}$ is the prior covariance matrix of $\xi_{j,k,0}$ given the initial information set $S^{-1}_{j,k}$. The above representation of the problem indicates that, given a quadratic objective and an initial Gaussian prior, the distribution of firms' belief about the state $\xi_{j,k,t}$ only matters through its conditional covariance matrices over time. We can then solve this problem using the same method as in Afrouzi and Yang (2019) and obtain the stationary pair $(\Sigma_{j,-1}, \Sigma_{j,0})$ such that given that $\Sigma_{j,0|-1} = \Sigma_{j,-1}$ then it is optimal for the agent to set $\Sigma_{j,t|t} = \Sigma_{j,0}, \forall t \geq 0$, and the initial prior is reproduced by $\Sigma_{j,0}$ in the sense that the law of motion for the covariance matrix of the state implies

$$\Sigma_{j,-1} = A_j \Sigma_{j,0} A'_j + \sigma_u^2 Q_j Q'_j \tag{J.22}$$

The interpretation of this procedure is that when the firm's ideal price follows the stationary process described above, if firm j,k's initial information set is such that $\mathbb{V}\text{ar}(\xi_{j,k,0}|S^{-1}_{j,k}) = \Sigma_{-1}$, the firm's optimal posterior beliefs about the state are also stationary over time and are given by Σ_0. One can interpret this similarly to Maćkowiak and Wiederholt (2009) as when the firm receives infinitely many signals at time

0 such that Σ_{-1} emerges as their initial prior.

A byproduct of this solution is that we obtain the shape of optimal signals that emerge under this information structure. In particular, we know that the firm receives a one-dimensional signal in this case because there is only one action taken at each period (see, e.g., Maćkowiak et al., 2018) or (Afrouzi & Yang, 2019, Lemma 1):

$$S_{j,k,t} = Y_j' \xi_{j,k,t} + e_{j,k,t} \tag{J.23}$$

where $Y_j \in \mathbb{R}^{n+m}$ is the loading of the signal and $e_{j,k,t} \sim N(0, \sigma_{j,e}^2)$ is the rational inattention noise of the firm that is orthogonal to the history of $(u_\tau, v_{j,-k,\tau})_{\tau \leq t}$.[64] Then, letting $\Lambda_j \equiv \Sigma_{j,-1} Y_j (Y_j' \Sigma_{j,-1} Y_j + \sigma_{j,e}^2)^{-1}$ to denote the Kalman gain that emerges under these signals for predicting $\xi_{j,k,t}$, we can write the firm's beliefs about the state as

$$\hat{\xi}_{j,k,t} \equiv \mathbb{E}[\xi_{j,k,t} | S_{j,k}^t] = A_j \hat{\xi}_{j,k,t-1} + \Lambda_j Y_j' (\xi_{j,k,t} - A_j \hat{\xi}_{j,k,t-1}) + \Lambda_j e_{j,k,t} \tag{J.24}$$

$$\Longrightarrow \hat{\xi}_{j,k,t} - \xi_{j,k,t} = (I - \Lambda_j Y_j') A_j (\hat{\xi}_{j,k,t-1} - \xi_{j,k,t-1}) + \Lambda_j e_{j,k,t} - (I - \Lambda_j Y_j') Q_j \epsilon_{j,k,t}, \quad \epsilon_{j,k,t} = (u_t, v_{j,-k,t})'$$

Now iterating this backward and replacing $\xi_{j,k,t} = \sum_{\tau=0}^{\infty} A_j^\tau Q_j \epsilon_{j,k,t-\tau}$, we obtain the projection of $\hat{\xi}_{j,k,t}$ on the history of shocks in $\epsilon_{j,k,t} = (u_t, v_{j,-k,t})'$ and the history of the firm's own rational inattention errors $e_{j,k,t}$:

$$\hat{\xi}_{j,k,t} = \sum_{\tau=0}^{\infty} \left[A_j^\tau - ((I - \Lambda_j Y_j') A_j)^\tau (I - \Lambda_j Y_j') \right] Q_j \epsilon_{j,k,t-\tau} + \sum_{\tau=0}^{\infty} ((I - \Lambda_j Y_j') A_j)^\tau \Lambda_j e_{j,k,t-\tau} \tag{J.25}$$

and, finally, since the optimal price of the firm is given by $p_{j,k,t} = H_j' \hat{\xi}_{j,k,t}$, we have:

$$p_{j,k,t} = \sum_{\tau=0}^{\infty} H_j' \left[A_j^\tau - ((I - \Lambda_j Y_j') A_j)^\tau (I - \Lambda_j Y_j') \right] Q_j \epsilon_{j,k,t-\tau} + \sum_{\tau=0}^{\infty} H_j' ((I - \Lambda_j Y_j') A_j)^\tau \Lambda_j e_{j,k,t-\tau} \tag{J.26}$$

which can be opened up as

$$p_{j,k,t} = \sum_{\tau=0}^{\infty} w_{j,u,\tau} u_{t-\tau} + \sum_{\tau=0}^{\infty} w_{j,v,\tau} v_{j,-k,t-\tau} + \sum_{\tau=0}^{\infty} w_{j,e,\tau} e_{j,k,t-\tau} \tag{J.27}$$

where

$$\begin{aligned} w_{j,u,\tau} &= H_j' \left[A_j^\tau - ((I - \Lambda_j Y_j') A_j)^\tau (I - \Lambda_j Y_j') \right] (Q_{j,u}', 0')' \\ w_{j,v,\tau} &= H_j' \left[A_j^\tau - ((I - \Lambda_j Y_j') A_j)^\tau (I - \Lambda_j Y_j') \right] (0', Q_{j,v}')' \\ w_{j,e,\tau} &= H_j' ((I - \Lambda_j Y_j') A_j)^\tau \Lambda_j \end{aligned} \tag{J.28}$$

Stage 2: The Fixed Point. Recall that, at Stage 1, we started by taking the stationary process of other firms' prices as given and solved for a single firm's optimal pricing strategy. In particular, we assumed

[64] Note that for the Kalman filtering problem, the scale of this signal is indeterminate; i.e., one can multiply the signal by any scalar without altering its optimality since such multiplication does not alter the signal to noise ratio; i.e., the Kalman gain vector Λ_j adjusts with the scale such that the scale of the signal is irrelevant for inference. The DRIPs.m package of Afrouzi and Yang (2019) sets this scale such that the the signal of the firm is its optimal price up to an additive constant (this makes the signal correspond to the equivalent of the recommendation strategies discussed in the static model).

that for $l \neq k$, $p_{j,l,t}$ has the following decomposition:

$$p_{j,l,t} = \psi_{j,u}(L)u_t + \psi_{j,v}(L)v_{j,l,t} \tag{J.29}$$

Now, in a symmetric equilibrium, the same $\psi_{j,u}(L)$ and $\psi_{j,v}(L)$ should also represent the pricing strategy of firm j,k. Therefore, we have two representations for j,k's price; one from the guess

$$p_{j,k,t} = \sum_{\tau=0}^{\infty} \psi_{j,u,\tau} u_{t-\tau} + \sum_{\tau=0}^{\infty} \psi_{j,v,\tau} v_{j,k,t-\tau} \tag{J.30}$$

and the other from the optimal pricing strategy derived under this guess:

$$p_{j,k,t} = \sum_{\tau=0}^{\infty} w_{u,\tau} u_{t-\tau} + \sum_{\tau=0}^{\infty} (w_{v,\tau} v_{j,-k,t-\tau} + w_{e,\tau} e_{j,k,t-\tau}) \tag{J.31}$$

It follows that in the symmetric stationary equilibrium, the following conditions should hold:

$$w_{j,u,\tau} = \psi_{j,u,\tau}, \forall \tau \geq 0 \tag{J.32}$$

$$\psi_{j,v,\tau} v_{j,k,t-\tau} = w_{j,v,\tau} v_{j,-k,t-\tau} + w_{j,e,t} e_{j,k,t-\tau}, \forall \tau \geq 0 \tag{J.33}$$

Note that the first equation already defines a part of the fixed point problem for the projection of prices on monetary shocks. However, to characterize we also need an updating rule for $\psi_{j,v,\tau}$, which can be obtained by making the following observations about $v_{j,k,t}, v_{j,-k,t}$ and $e_{j,k,t}$. Since $e_{j,k,t} \perp v_{j,-k,t}$ we can take the variance of both sides of the second equation above to obtain:

$$\psi^2_{j,v,\tau} \mathbb{V}\text{ar}(v_{j,k,t-\tau}) = w^2_{j,v,\tau} \sigma^2_u + w^2_{j,e,\tau} \sigma^2_{e,j} \tag{J.34}$$

where we have already plugged in the normalization that $\mathbb{V}\text{ar}(v_{j,-k,t}) = \sigma^2_u$. Second, since the distribution of $(v_{j,l,t})_{l \in K_j}$ should be symmetric in the symmetric stationary equilibrium, we obtain the following condition by taking the covariance of both sides with $v_{j,-k,t}$:

$$\psi_{j,v,\tau} \mathbb{C}\text{ov}(v_{j,k,t-\tau}, v_{j,-k,t-\tau}) = w_{j,v,\tau} \sigma^2_u \tag{J.35}$$

Finally, by symmetry and our previous normalization of $\mathbb{V}\text{ar}(v_{j,-k,t}) = \sigma^2_u$:[65]

$$\sigma^2_u = \mathbb{V}\text{ar}(v_{j,-k,t-\tau}) = \mathbb{V}\text{ar}\left(\frac{1}{K_j - 1} \sum_{l \neq k} v_{j,l,t-\tau}\right) = \frac{1}{K_j - 1} \mathbb{V}\text{ar}(v_{j,k,t-\tau}) + \frac{K_j - 2}{K_j - 1} \mathbb{C}\text{ov}(v_{j,k,t-\tau}, v_{j,-k,t-\tau}) \tag{J.36}$$

Multiplying both sides of this last equation by $\psi^2_{j,v,\tau}$ and substituting the first two equations above, we obtain:

$$\psi^2_{j,v,t} \sigma^2_u = \frac{1}{K_j - 1} (w^2_{j,v,\tau} \sigma^2_u + w^2_{j,e,\tau} \sigma^2_{e,j}) + \frac{K_j - 2}{K_j - 1} \psi_{j,v,\tau} w_{j,v,\tau} \sigma^2_u \tag{J.37}$$

which is a quadratic equation in $\psi_{j,v,t}$ and has only one positive root given $w_{j,v,\tau}$ and $w_{j,e,\tau}$.

Therefore, we have a the following mapping between $(\psi_{j,u,\tau}, \psi_{j,v,\tau})$ and $(w_{j,u,\tau}, w_{j,v,\tau}, w_{j,e,\tau})$:

$$\psi_{j,u,\tau} = w_{j,u,\tau}, \forall \tau \geq 0 \tag{J.38}$$

[65] In deriving this equation, I have used $\mathbb{V}\text{ar}(v_{j,k,t}) = var(v_{j,l,t})$ and $\mathbb{C}\text{ov}(v_{j,k,t}, v_{j,-k,t}) = \mathbb{C}\text{ov}(v_{j,k,t}, v_{j,l,t})$ for any $l \neq k$.

$$\psi_{j,v,\tau} = \frac{1}{2}\left(\frac{K_j-2}{K_j-1}w_{j,v,\tau} + \sqrt{\frac{(K_j)^2}{(K_j-1)^2}w_{j,v,\tau}^2 + \frac{4}{K_j-1}w_{j,e,\tau}^2(\frac{\sigma_{e,j}}{\sigma_u})^2}\right), \forall \tau \geq 0 \tag{J.39}$$

Finally, recall that in the definition of a symmetric equilibrium (Definition 1) we are looking within strategies in which all sectors with the same number of competitors have the same pricing strategies; thus, the index j on the lag polynomials $\psi_{j,u}(L)$ and $\psi_{j,v}(L)$ emphasizes that these coefficients can vary for different values of K_j but not with j directly. As a result in the rest of this section, we will often use notation $\psi_{K,u,\tau}$ and $\psi_{K,v,\tau}$ to emphasize that these coefficients are the same for all firms in a sector with K competitors. The solution of the model then boils down to characterizing the coefficients of the lag polynomials for every K in the support of the distribution of the number of competitors $\mathcal{K}$, $(\psi_{K,u,\tau}, \psi_{K,v,\tau})_{K \in \text{Supp}(\mathcal{K})}^{\tau \geq 0}$, such that the fixed point condition in Equations (J.38) and (J.39) holds for all $K \in \mathbb{N}$ and $\tau \geq 0$.

Once we have these coefficients, we can then construct the impulse response of prices and output to monetary shocks for sectors with K competitors using the symmetry of responses within such sectors, and then construct the impulse response of aggregate price and output by weighting the responses of sectors with K competitors by their shares in the distribution of number of competitors, $\mathcal{K}$.[66] Formally, letting s_K denote the share of firms with K competitors in the distribution of the number of competitors, we can write the aggregate price as:

$$p_t = \sum_{j \in J} \frac{1}{JK_j} \sum_{k \in K_j} p_{j,k,t} = \sum_{j \in J} \frac{1}{JK_j} \sum_{k \in K_j} (\psi_{K_j,u}(L)u_t + \psi_{K,v}(L)v_{j,-k,t}) \tag{J.40}$$

Now, noting that mistakes are independent across sectors and J is large (because firms only pay attention to mistakes of firms within their own sector but across sectors), the term involving mistakes washes out in the aggregate prices and we have:

$$p_t = \sum_{j \in J} \frac{1}{JK_j} \sum_{k \in K_j} \psi_{K_j,u}(L)u_t = \sum_{K \in \text{Supp}(\mathcal{K})} s_K \psi_{K,u}(L)u_t = \psi_{p,u}(L)u_t, \qquad \psi_{p,u}(L) \equiv \sum_{K \in \text{Supp}(\mathcal{K})} s_K \psi_{K,u}(L) \tag{J.41}$$

It then follows that output is given by

$$y_t = q_t - p_t = (\psi_{q,u}(L) - \psi_{p,u}(L))u_t \tag{J.42}$$

J.3. Integrated MA State Space Representation and Solution Algorithms

In this section, I discuss the first algorithm that I use to find the fixed points for the coefficients of the lag polynomials $(\psi_{K,u,\tau}, \psi_{K,v,\tau})_{K \in \text{Supp}(\mathcal{K})}^{\tau \geq 0}$, which is based on an integrated MA truncation algorithm.

We derived the fixed point conditions in Equations (J.38) and (J.39) by relying on an approximation

[66]First note that in the benchmark model, $p_{j,t} + y_{j,t} = q_t$ since sectoral goods are neither complements nor substitutes (Cobb-Douglas preferences imply that the expenditure share of household on sector j is constant so in log-deviations total nominal demand moves one to one with sectoral nominal demand of j). Thus, $y_{j,t}$ can be constructed as $y_{j,t} = q_t - p_{j,t}$ once we have $p_{j,t}$. Moreover, note that in the symmetric steady-state around which we have linearized the economy, the total expenditure share of the household on sectors with K competitors is simply the share of such firms in the distribution of the number of competitors, $\mathcal{K}$. So the aggregate price p_t, which is the expenditure share weighted price across all sectors, can be calculated by summing up the responses of sectors with K competitors weighted by their shares in $\mathcal{K}$. Aggregate output is then given by the difference between the nominal GDP, q_t, and the aggregate price, p_t, as $q_t = p_t + y_t$.

of the process for the ideal price of firms in Equation (J.15) with a Markov state space representation in Equation (J.16). Using Equations (J.13) and (J.15) we can write the ideal price of firm j,k (where j is an industry with K competitors) as:

$$p^*_{j,k,t} = \underbrace{\sum_{\tau=0}^{\infty}((1-\alpha_K)\psi_{q,u,\tau}+\alpha_K\psi_{K,u,\tau})u_{t-\tau}}_{\equiv x^u_{K,t}} + \underbrace{\alpha_j\sum_{\tau=0}^{\infty}\psi_{K,v,\tau}v_{j,-k,t-\tau}}_{\equiv x^v_{j,k,t}} \tag{J.43}$$

Thus, the coefficients $(\psi_{q,u,\tau},\psi_{K,u,\tau},\psi_{K,v,\tau})_{\tau=0}^{\infty}$ constitute the MA(∞) representation of the ideal price with respect to shocks to nominal GDP growth and the average mistakes of other firms in j's sector.

Now, if prices were stationary, then these sequences would have been square summable (often denoted as being in ℓ^2), and thus their coefficients must have converged to zero as $\tau \to \infty$. Hence, one could have approximated these coefficients arbitrarily well with an MA(T) process for large T. However, since nominal GDP is assumed to have a unit root, this is not a proper approximation by itself as $\psi_{q,u,\tau}$ is no longer square summable. Moreover, prices themselves also inherit the unit root from q_t.[67] Therefore, neither $\psi_{q,u,\tau}$ nor $\psi_{K,u,\tau}$ converge to 0 as $\tau\to\infty$. Thus, we need an approach for approximating these polynomials that takes these unit roots into account. As for the mistakes polynomial coefficients, $\psi_{K,v,\tau}$, these are square summable because they are only a function of the current and past signal noises ($e_{j,k,t}: k\in K_j$) of firms in sector j. Since signal noises of any firm are i.i.d. over time, and firms put less and less weight on their past signals to partly take their newer signals into account, we can truncate the MA(∞) representation of these arbitrarily well with an MA(T) process.

The idea behind finding a state space representation for the Markov process that respects the unit roots of q_t and $p_{j,-k,t}$ is to use an integrated MA truncation by defining a random walk process in terms of monetary shocks:

$$\tilde{u}_t = \frac{u_t}{1-L} = \sum_{\tau=0}^{\infty} u_{t-\tau} \tag{J.44}$$

We can then re-write the ideal price of firm j,k in terms of $\tilde{u}_t$:

$$p^*_{j,k,t} = (1-\alpha_K)\psi_{q,u}(L)u_t + \alpha_K\psi_{K,u}(L)u_t + \alpha_K\psi_{K,v}(L)v_{j,-k,t} \tag{J.45}$$

$$= (1-\alpha_K)\Delta\psi_{q,u}(L)\tilde{u}_t + \alpha_K\Delta\psi_{K,u}(L)\tilde{u}_t + \alpha_K\psi_{K,v}(L)v_{j,-k,t} \tag{J.46}$$

where we have defined $\Delta\psi_{q,u}(L)\equiv(1-L)\psi_{q,u}(L)$ and $\Delta\psi_{K,u}(L)\equiv(1-L)\psi_{K,u}(L)$. It is easy to verify that the coefficients of these two polynomials are square summable. To see why, we can make the observation that they correspond to the IRFs of the growth in nominal GDP and prices of other firms:

$$\Delta q_t = (1-L)q_t = (1-L)\psi_{q,u}(L)u_t = \Delta\psi_{q,u}(L)u_t \tag{J.47}$$

$$\Delta p_{j,-k,t} = (1-L)p_{j,-k,t} = (1-L)\psi_{K,u}(L)u_t = \Delta\psi_{K,u}(L)u_t \tag{J.48}$$

Now, since both q_t and $p_{j,-k,t}$ have exactly one unit root, their difference, Δq_t and $\Delta p_{j,-k,t}$, are stationary

[67] To see why, note that if prices do not have exactly one unit root, then any firm's losses, $((1-\alpha_j)q_t+\alpha_j p_{j,-k,t}-p_{j,k,t})^2$, would grow unboundedly so optimal information acquisition around a steady state with bounded profit loss implies that prices should have exactly one unit root

and their IRFs are square summable. This implies that the coefficients $\Delta\psi_{q,u,\tau}$ and $\Delta\psi_{K,u,\tau}$ converge to 0 for large τ. So for a given level of tolerance, we can find T_u and T_v such that for $\tau \geq T_u$ and $\tau \geq T_v$, $\Delta\psi_{q,u,\tau}$ and $\Delta\psi_{K,u,\tau}$ are approximately zero under that tolerance level.[68]

Now, consider the following state space representation. Fixing T_u and T_v as above, let $U^t_{j,k} \equiv (\tilde{u}_t,\ldots,\tilde{u}_{t-T_u},v_{j,-k,t},\ldots,v_{j,-k,T_v})' \in \mathbb{R}^{T_u+T_v}$. Then, we can write the law of motion for $U^t_{j,k}$ as:

$$\underbrace{\begin{bmatrix} \tilde{u}_t \\ \tilde{u}_{t-1} \\ \tilde{u}_{t-2} \\ \vdots \\ \tilde{u}_{t-T_u} \end{bmatrix}}_{\equiv \tilde{U}^t} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}}_{\equiv A_u} \underbrace{\begin{bmatrix} \tilde{u}_{t-1} \\ \tilde{u}_{t-2} \\ \tilde{u}_{t-3} \\ \vdots \\ \tilde{u}_{t-T_u-1} \end{bmatrix}}_{=\tilde{U}^{t-1}} + \sigma_u \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\equiv Q_u} u_t \tag{J.49}$$

so that the first row gives $\tilde{u}_t = \tilde{u}_{t-1} + u_t$ and the rest of the rows give $\tilde{u}_{t-h} = \tilde{u}_{t-h}, \forall h \geq 1$. Moreover, note that

$$\underbrace{\begin{bmatrix} v_{j,-k,t} \\ v_{j,-k,t-1} \\ v_{j,-k,t-2} \\ \vdots \\ v_{j,-k,t-T_v} \end{bmatrix}}_{\equiv V^t_{j,-k}} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}}_{\equiv A_v} \underbrace{\begin{bmatrix} v_{j,-k,t-1} \\ v_{j,-k,t-2} \\ v_{j,-k,t-3} \\ \vdots \\ v_{j,-k,t-T_v-1} \end{bmatrix}}_{\equiv V^t_{j,-k}} + \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\equiv Q_v} v_{j,-k,t} \tag{J.50}$$

so that each row gives $v_{j,-k,t-h} = v_{j,-k,t-h}$. Given these laws of motion, we then have the following augmented state space representation for $U^t_{j,k}$:

$$U^t_{j,k} = \begin{bmatrix} \tilde{U}^t \\ V^t_{j,-k} \end{bmatrix} = \underbrace{\begin{bmatrix} A_u & 0_{T_u \times T_v} \\ 0_{T_v \times T_u} & A_v \end{bmatrix}}_{\equiv A} U^{t-1}_{j,k} + \underbrace{\begin{bmatrix} Q_u & 0_{T_u \times 1} \\ 0_{T_v \times 1} & Q_v \end{bmatrix}}_{\equiv Q} \begin{bmatrix} u_t \\ v_{j,-k,t} \end{bmatrix} \tag{J.51}$$

We then approximate the ideal price of firm j,k as:

$$p^*_{j,k,t} \approx \hat{p}^*_{j,k,t} \equiv \sum_{\tau=0}^{T_u} ((1-\alpha_{K_j})\Delta\psi_{q,u,\tau} + \alpha_{K_j}\Delta\psi_{K_j,u})\tilde{u}_{t-\tau} + \alpha_j \sum_{\tau=0}^{T_v} \psi^{\tau}_{K_j,v} v_{j,-k,t-\tau} = H'_{K_j} U^t_{j,k} \tag{J.52}$$

[68]In the code for solving and calibrating the benchmark model, I set that $T_u = 60$ and $T_v = 30$ (i.e. implicitly assuming that inflation and output responses to monetary converge to zero within 60 quarters (15 years). These are then confirmed in the implied IRFs of the calibrated model, where both these responses converge to zero within 12 to 16 quarters. See, e.g., Figure A.5). Thus, these values are large enough that, but small enough that the state space representation is not too large (90×90). Additionally, I solve the model also with $T_u = 40$ and $T_v = 20$, obtaining identical results. In that sense, $T_u = 60$ and $T_v = 30$ are conservative choices that allow for longer truncations, but these turn out to be unnecessary. Beyond the benchmark model, I set $T_u = 40$ and $T_v = 20$ across some of the robustness exercises in Appendices L and M.

where

$$H_{K_j} = (1-\alpha_{K_j}) \begin{bmatrix} \Delta\psi_{q,u,0} \\ \vdots \\ \Delta\psi_{q,u,T_u} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \alpha_K \begin{bmatrix} \Delta\psi_{K_j,u,0} \\ \vdots \\ \Delta\psi_{K_j,u,T_u} \\ \psi_{K_j,v,0} \\ \vdots \\ \psi_{K_j,mv,T_v} \end{bmatrix} \tag{J.53}$$

Note that this integrated MA truncation is only inaccurate because it implicitly assumes that $\Delta\psi_{K,u,\tau} = 0, \forall \tau \geq T_u + 1$ and $\psi_{K,v,\tau} = 0, \forall \tau \geq T_v + 1$. But since these coefficients are square summable, for large enough T_u and T_v, we know that $(\Delta\psi_{K,u,\tau})_{\tau=T_u}^{\infty}, (\psi_{K,v,\tau})_{\tau=T_v}^{\infty}$ are arbitrarily small and close to zero.

Moreover, one subtlety here is that by differencing out the lag polynomials, we have pushed the unit root of the process to the state space itself (as captured by the fact that the matrix A now has an eigenvalue on the unit circle). Accordingly, the unconditional covariance matrix of $U_{j,k}^t$ is now unbounded, which might undermine the accuracy of this approximation. However, the key is to recognize that this approximation only needs to be accurate from the perspective of a firm that takes the process of its ideal price as given and solves its rational inattention problem. Such a firm does not evaluate the variance-covariance matrix of the $U_{j,k}^t$ unconditionally, but rather conditional on their information set. It is then straightforward to see that a covariance matrix that is growing unboundedly cannot be an optimal choice for firms in equilibrium, as it would mean that their losses from mispricing grow unboundedly with time. However, since the cost of attention is only logarithmic in the determinant of this covariance matrix, as uncertainty gets larger, the marginal cost of reducing that uncertainty becomes arbitrarily smaller for the firm. So it must be that the equilibrium covariance matrix for $U_{j,k}^t$ is bounded and the approximation above is accurate.[69]

So the fixed point problem boils down to finding the finite sequences $(\Delta\psi_{K,u,\tau})_{\tau=0}^{T_u}, (\psi_{K,v,\tau})_{\tau=0}^{T_v}$, which we can solve using the following algorithm:

Algorithm 1 (Solving the Model with Integrated MA Approximation)**.** For a given $K \in \text{Supp}(\mathcal{K})$:

1. Start with a guess for $(\Delta\psi_{K,u,\tau})_{\tau=0}^{T_u}, (\psi_{K,v,\tau})_{\tau=0}^{T_v}$ (at iteration 0, set them equal to their values under rational expectations with full information: $\Delta\psi_{K,u,\tau} = \rho^\tau$ and $\psi_{K,v,\tau} = 0$).
2. Form matrices A, Q, and H_K using Equations (J.51) and (J.53). Solve the rational inattention problem of the firm in Equation (J.19) using the method in Afrouzi and Yang (2019).
3. Using the solution to the dynamic rational inattention problem, find the stationary pair $(\Sigma_{j,-1}, \Sigma_{j,0})$

[69] See also the discussion in Section 4.4 of Maćkowiak et al. (2018). In that paper, the dual rational inattention problem is studied where the capacity of processing information κ is exogenous. Therefore, an additional condition is that κ needs to be large enough. Here, since we are considering the problem where the firm also chooses κ endogenously, it is then implied that the equilibrium κ would be such that equilibrium beliefs are finite; otherwise, firms are making unboundedly large losses in profits from mispricing, which cannot be an optimal choice of κ. For the simplest working example, see also the simple pricing model in Afrouzi and Yang (2019) which assume that the nominal GDP process is a random walk, and shows that the equilibrium beliefs of firms are bounded above by a reservation uncertainty level. More generally, the proofs provided in that paper for the characterization of stationary covariance matrices go through when A has an eigenvalue on the unit circle.

and the implied Y_j, Λ_j and $\sigma^2_{e,j}$ to construct the time-invariant IRFs of a firm's optimal price with respect to shocks in Equations (J.27) and (J.28).

4. Use Equations (J.38) and (J.39) to derive the implied $(\psi_{K,u,\tau})_{\tau=0}^{T_u}$,$(\psi_{K,v,\tau})_{\tau=0}^{T_v}$.
5. Update the guess for $(\Delta\psi_{K,u,\tau})_{\tau=0}^{T_u}$,$(\psi_{K,v,\tau})_{\tau=0}^{T_v}$ using the implied values from Step 4 and repeat the process in Steps 1 to 5 until convergence.
6. Once the fixed point is found for all $K \in \text{Supp}(\mathcal{K})$, construct the aggregate IRFs of prices and output using Equations (J.41) and (J.42).

Implementation in the Replication Package. To conclude this section, I briefly discuss how Algorithm 1 is implemented in the replication package of the project. The code in the replication package is automated to produce all model-based results through the single file `./main.m`. In particular, lines 19 to 32 of this file are switches that take values `'Y'` (yes) or `'N'` (no) that determine whether the user wants to replicate a particular result. Each switch, when set to `'Y'`, then calls a particular part of the `'main.m'` file that sets up the parameters, solution method options, as well as simulation and calibration options when required and dispatches the proper internal functions to solve, simulate and calibrate the respective model.

Specifically, the `replicate.Calibrate_Benchmark` switch replicates the solution, simulation, and calibration of the benchmark model using Algorithm 1. The two files that implement the solution of the model using this algorithm are `./codes/matlab/solve_models_int_ma.m` as well as the accompanying file `./codes/matlab/solve_model_int_ma.m`. To briefly describe what each of these functions does, the file `solve_model_int_ma.m` solves the model for a given set of parameters using Algorithm 1. This file is exclusively called by `solve_model_int_ma.m`, which takes an array of values of $K \in \text{Supp}(\mathcal{K})$ as well as other parameters (e.g., multiple values of ω) and dispatches multiple instances of the first file for parallel computation in order to solve several models simultaneously. A key observation about Algorithm 1 is that it requires simultaneous convergence of the `DRIPs.m` algorithm of Afrouzi and Yang (2019) (this happens in Step 3 of the algorithm when we find the matrices Y_j,$\Sigma_{j,-1}$,$\Sigma_{j,0}$,$\sigma^2_{e,j}$ given a guess for the coefficients of lag polynomials) as well as the convergence of the lag polynomials themselves (which happens in Step 5). It turns out that doing these sequentially is computationally expensive, so in the `solve_model_int_ma.m`, I further parallelize the convergence of these objects using the following procedure: for every guess of the coefficients of the lag polynomials, I only do a small number of iterations in the `DRIPs.m` package (10 iterations) but augment the convergence error of the rational inattention problem to the convergence of Algorithm 1 (this is returned automatically by the `DRIPs.m` package as `ri.ss.err`). Then, I compute the implied coefficients for the lag polynomials in an inner loop. The algorithm confirms convergence when *all* convergence errors are small; i.e., both `ri.ss.err` and the difference between the coefficients of the lag polynomials in two consecutive iterations are *jointly* smaller than a given tolerance level. With this approach, the algorithm avoids solving the dynamic rational inattention problem fully for every wrong guess of the coefficients for the lag polynomials, but once everything converges, it is implied that the rational inattention problem is also solved properly. This updated algorithm is very fast in solving the problem, which is necessary for solving the model with high precision for a large number of values of $K\in\text{Supp}(\mathcal{K})$ and ω's required for calibration.

Once the model is solved for a given ω and all $K\in\text{Supp}(\mathcal{K})$, `solve_models_int_ma.m` returns a structure containing all the impulse response functions of the model for different values of K as well as the IRFs of aggregate inflation and aggregate output. The file `./codes/matlab/simulate.m` then simulates the model for a large number of firms and computes their forecasts and nowcasts of inflation over a time series of length T, burns some initial periods, and generates two cross-sections of firms' expectations, similarly apart in time according to the survey evidence I use to calibrate ω. It then adds this simulated dataset to the solution structure of the model and returns this augmented structure. The function `./codes/matlab/calib_eval.m` then runs the regression for the calibration of ω in the simulated data and calculates the model implied moment. It then computes the distance between this simulated moment and its equivalent from the data in Table 3 and returns the quadratic difference between the two. Finally, the function `./codes/calibrate.m` automates the procedure of minimizing this calibration loss function to find the value of ω that minimizes the distance between the data simulated moment and its empirical counterpart using a Nelder-Mead optimization algorithm. To confirm identification, I then resolve the model on a grid of ω's around the optimized value of ω and plot the simulated and empirical moments as a function of ω in Figure A.3 to show that (1) the empirical moment is informative of ω in the model and (2) the value of ω returned by the Nelder-Mead algorithm matches this moment well.

Finally, it is worth noting that in order to standardize the initial guesses for ω across different models and parameter values, I use the following change of variables in the code. Recall the dynamic rational inattention problem of a firm in Equation (J.19):

$$\max_{\{\Sigma_{j,t|t}\succsim 0\}_{t\geq 0}} -\frac{1}{2}\sum_{t=0}^{\infty}\left[B_j tr(H_jH_j'\Sigma_{j,t|t})+\omega\ln\left(\frac{|\Sigma_{j,t|t-1}|}{|\Sigma_{j,t|t}|}\right)\right] \tag{J.54}$$

$$s.t.\quad \Sigma_{j,t+1|t}=A_j\Sigma_{j,t|t}A_j'+\sigma_u^2Q_jQ_j',\quad \forall t\geq 0 \tag{J.55}$$

$$\Sigma_{j,t|t-1}-\Sigma_{j,t|t}\succsim 0\quad \forall t\geq 0,\qquad \Sigma_{0|-1}\equiv\mathbb{V}\text{ar}(\xi_{j,k,0}|S_{j,k}^{-1}) \tag{J.56}$$

When coding this problem, I utilize the following change of variables: $\tilde{\omega}\equiv\frac{\omega}{B_\infty\sigma_u^2}$ where $B_\infty=\lim_{K_j\to\infty}B_j$ from Equation (28), $\Sigma_{j,t|t}^n\equiv\Sigma_{j,t|t}/\sigma_u^2$, and $\Sigma_{j,t|t-1}^n\equiv\Sigma_{j,t|t-1}/\sigma_u^2$. It then follows that the problem can be written as

$$\max_{\{\Sigma_{j,t|t}^n\succsim 0\}_{t\geq 0}} -\frac{B_j\sigma_u^2}{2}\sum_{t=0}^{\infty}\left[tr(H_jH_j'\Sigma_{j,t|t}^n)+\tilde{\omega}\frac{B_\infty}{B_j}\ln\left(\frac{|\Sigma_{j,t|t-1}^n|}{|\Sigma_{j,t|t}^n|}\right)\right] \tag{J.57}$$

$$s.t.\quad \Sigma_{j,t+1|t}^n=A_j\Sigma_{j,t|t}^nA_j'+Q_jQ_j',\quad \forall t\geq 0 \tag{J.58}$$

$$\Sigma_{j,t|t-1}^n-\Sigma_{j,t|t}^n\succsim 0\quad \forall t\geq 0,\qquad \Sigma_{0|-1}^n\equiv\mathbb{V}\text{ar}(\xi_{j,k,0}|S_{j,k}^{-1})/\sigma_u^2 \tag{J.59}$$

This is a normalized version of the problem that harmonizes the solution across different parameter values for optimization over $\tilde{\omega}$. Once I have solved this problem for a given set of parameters including $\tilde{\omega}$, I then recover $\omega=\sigma_u^2B_\infty\tilde{\omega}$, $\Sigma_{t|t-1}=\sigma_u^2\Sigma_{t|t-1}^n$ and $\Sigma_{t|t}=\sigma_u^2\Sigma_{t|t}^n$.

J.4. Robustness: ARMA Approximation

In this section, I discuss the robustness of the integrated MA truncation approach to an ARMA approximation of the state space representation, as in Maćkowiak et al. (2018). The conclusion is that for the values of T_u and T_v that I use in the paper, the results of the integrated MA truncation are numerically identical to an ARMA approximation.

To briefly summarize the robustness of results to this alternative approximation, the maximum distance between the output and inflation IRFs to a 100 basis points monetary shock, averaged across all values of K in the model, are only 0.37 and 0.13 basis points, respectively.[70] Figure J.1 also shows the IRFs of output and inflation for three different values of K across two methods and confirms visually that these impulse responses obtained from the two algorithms appear identical.

Furthermore, Tables J.1 to J.3 present the analogs of Tables 4 to 6 in the main text, under this section's ARMA approximation. The results are numerically identical to the results of the integrated MA truncation in Tables 4 to 6. Therefore, the results of the paper are robust to this alternative solution method.

The rest of this section describes the details of the ARMA approximation, presents an algorithm for solving the model under this approximation, and concludes with a discussion of how this algorithm is implemented in the replication package of the paper.

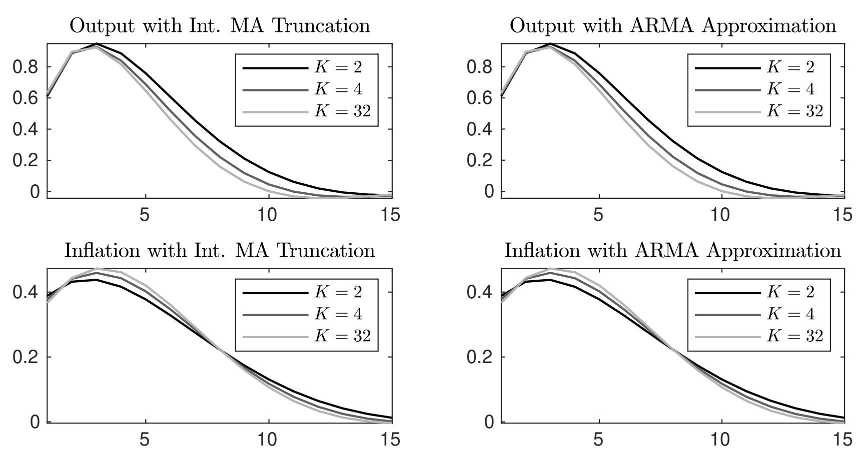


Figure J.1: IRFs under Integrated MA Truncation and ARMA Approximation Solution Methods

Notes: For different values of K under the benchmark calibration of the model, the figure plots the impulse response of output (top panel) and inflation (bottom panel) to a 1% expansionary monetary policy shock under the integrate MA truncation of Algorithm 1 (left panel) and the ARMA approximation of Algorithm 2 (right panel).

[70] This is the average number across sectors with different K. Alternatively, instead of calculating the mean of these maximum differences across different values of K, we can measure the maximum value of these maximum differences, which is the total upper bound for the differences between IRFs generated by the two algorithms. For a 100 basis points monetary shock, this max-max distance is 0.26 basis points for inflation and 0.70 basis points for output.

Table J.1: Output and Monetary Non-Neutrality Across Models (Robustness to ARMA Approximation)

Model		*Variance*		*Persistence*	
		var(Y) $^{\times 10^4}$	*amp. factor*	*half-life* ^{qtrs}	*amp. factor*
		(1)	*(2)*	*(3)*	*(4)*
Monopolistic Competition		3.17	1.00	3.41	1.00
Benchmark	$K \sim \hat{\mathcal{K}}$	4.07	1.28	3.71	1.09
2-Competitors	$K = 2$	4.69	1.48	4.12	1.21
4-Competitors	$K = 4$	4.14	1.30	3.79	1.11
8-Competitors	$K = 8$	4.00	1.26	3.62	1.06
16-Competitors	$K = 16$	3.94	1.24	3.58	1.05
32-Competitors	$K = 32$	3.92	1.23	3.56	1.04
∞-Competitors	$K \to \infty$	3.89	1.23	3.55	1.04

Notes: The table presents the analog of Table 4 (which was computed using the integrated MA approximation) for monetary non-neutrality across models with different numbers of competitors under the ARMA approximation of Appendix J.4. Results are numerically identical across the two tables with minute third digits differences in Columns (3) and (4).

Table J.2: Inflation Across Models (Robustness to ARMA Approximation)

Model		*Variance*		*Persistence*	
		var(π) $^{\times 10^4}$	*damp. factor*	*half-life* ^{qtrs}	*amp. factor*
		(1)	*(2)*	*(3)*	*(4)*
Monopolistic Competition		1.47	1.00	4.42	1.00
Benchmark	$K \sim \hat{\mathcal{K}}$	1.37	0.94	4.66	1.05
2-Competitors	$K = 2$	1.28	0.87	4.83	1.09
4-Competitors	$K = 4$	1.36	0.93	4.68	1.06
8-Competitors	$K = 8$	1.39	0.95	4.64	1.05
16-Competitors	$K = 16$	1.40	0.95	4.62	1.05
32-Competitors	$K = 32$	1.41	0.96	4.62	1.05
∞-Competitors	$K \to \infty$	1.41	0.96	4.61	1.04

Notes: The table presents the analog of Table 5 (which was computed using the integrated MA approximation) for inflation response across models with different numbers of competitors under the ARMA approximation of Appendix J.4. Results are numerically identical across the two tables.

Table J.3: Strategic Inattention vs. Real Rigidities (Robustness to ARMA Approximation)

	Percentage change in variance of	
	output	*inflation*
	(1)	*(2)*
Total Change (percent)	18.6	-9.7
Due to Str. Inattention (ppt)	78.6	-19.8
Due to Real Rigidities (ppt)	-60.0	10.1

Notes: The table presents the analog of Table 6 (which was computed using the integrated MA approximation) for the decomposition of the effects of the strategic inattention and real rigidity channels for the change in volatility of output (monetary non-neutrality) and inflation conditional on monetary shocks under the ARMA approximation of Appendix J.4.

Solving the Model using an ARMA Approximation. To briefly discuss the main concern that leads to this exercise, note that if inflation or output have very persistent responses, the integrated MA truncation approach could possibly fail to capture these persistent effects. A solution for this, as proposed in Maćkowiak et al. (2018), is to approximate the implied IRFs of the inflation and output process with an ARMA process rather than an MA process to allow for such potential persistent effects to be captured by the AR coefficients.

To do so, I construct the following alternative state space representation. First, recall from Equation (J.43) that:

$$p^*_{j,k,t} = x^u_{K_j,t} + x^v_{j,k,t} \tag{J.60}$$

where $x^u_{K_j,t}$ and $x^v_{j,k,t}$ are the projection of firm j,k's ideal price on the monetary and mistake shocks, respectively. The main issue, as discussed in the previous section, is that $x^u_{K_j,t}$ has a unit root and we want to have a state space representation that approximates this process properly. To do so, we start by doing an ARMA approximation of $\Delta x^u_{K_j,t}$. Since $x^u_{K_j,t}$ has exactly one unit root (as discussed above), $\Delta x^u_{K_j,t}$ is a stationary process and can be approximated arbitrarily well with an ARMA(p,q) process (Han et al., 2022; Maćkowiak et al., 2018):

$$\Delta x^u_{K_j,t} \approx \sum_{i=1}^{p} \phi_{j,i} \Delta x^u_{K_j,t-i} + \sum_{i=0}^{q} \theta_{j,i} u_{t-i} \iff \Delta x^u_{K_j,t} \approx (1 - \sum_{i=1}^{p} \phi_{j,i} L^i)^{-1} \sum_{i=0}^{q} \theta_{j,i} L^i u_t \tag{J.61}$$

To implement this into Algorithm 1, I use the state-space representation for ARMA(p,q) processes proposed by (Hamilton, 1994, Ch. 13, p. 375, Equations 13.1.21-23) whose dimension is given by $r = \max\{p, q+1\}$. In particular, (Hamilton, 1994, Ch. 13) defines:

$$\xi_{j,t} \equiv (1 - \sum_{i=1}^{p} \phi_{j,i} L^i)^{-1} u_t \implies (1 - \sum_{i=1}^{p} \phi_{j,i} L^i) \xi_{j,t} = u_t \tag{J.62}$$

which implies that

$$\Delta x^u_{K_j,t} = \sum_{i=0}^{q} \theta_{j,i} L^i \xi_{j,t} = \sum_{i=0}^{q} \theta_{j,i} \xi_{j,t-i} \tag{J.63}$$

Now, in matrix form, following note that $\xi_{j,t}$ can be represented by defining:

$$\vec{\xi}_{j,t} \equiv \begin{bmatrix} \phi_{j,1} & \phi_{j,2} & \phi_{j,3} & \dots & \phi_{j,r} \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \vec{\xi}_{j,t-1} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u_t \tag{J.64}$$

and noticing that

$$\Delta x^u_{K_j,t} = (\theta_{j,0}, \theta_{j,1}, \dots, \theta_{j,r}) \cdot \vec{\xi}_{j,t} \tag{J.65}$$

where we interpret $\phi_{j,i} = 0$ for $i > p$ and $\theta_{j,i} = 0$ for $i > q$.

Finally, note that the rational inattention problem requires a state space representation for $x^u_{j,t}$ and not $\Delta x^u_{j,t}$, but note that once we have a ARMA(p,q) state space representation for the former, we can derive an ARIMA(p,1,q) state space representation for the latter using the following modification:

$$(1-\sum_{i=1}^{r}\phi_{j,i}L^i)\Delta x^u_{j,t}=\sum_{i=0}^{q}\theta_{j,i}u_t \Longrightarrow (1-\sum_{i=1}^{r}\phi_{j,i}L^i)(1-L)x^u_{j,t}=\sum_{i=0}^{q}\theta_{j,i}L^iu_t \tag{J.66}$$

$$\Longrightarrow (1-(1+\phi_{j,1})L-\sum_{i=2}^{p-1}(\phi_{j,i}-\phi_{j,i-1})L^i+\phi_{j,i}L^{i+1})x^u_{j,t}=\sum_{i=0}^{q}\theta_{j,i}L^iu_t \tag{J.67}$$

or, in matrix form, if we define $\Xi^u_{j,t}$ as:

$$\Xi^u_{j,t}\equiv \underbrace{\begin{bmatrix} 1+\phi_{j,1} & \phi_{j,2}-\phi_{j,1} & \dots & \phi_{j,r}-\phi_{j,r-1} & -\phi_{j,r} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}}_{A_{j,u}} \Xi_{j,t-1}+\underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{Q_u} u_t \tag{J.68}$$

where I have indexed A_j with j to emphasize that the coefficients $\phi_{j,i}$ depend on the lag polynomials of sector j's prices and will vary with K_j. Moreover, once we have this state space representation, we can recover $x^u_{K_j,t}$ as:

$$x^u_{K_j,t}=\underbrace{(\theta_0,\theta_1,\dots,\theta_r)}_{\equiv H_{j,u}}\cdot \Xi^u_{j,t} \tag{J.69}$$

where $\theta_{j,i}\equiv 0$ for $i>q$. This also implies once we adopt the (Hamilton, 1994, Ch. 13) state space representation for $\Delta x^u_{K_j,t}$ and fix the dimension of the state vector $\Xi_{j,t}$, which is $\max\{p+1,q+1\}$, we can set $r=p=q$ to take the most advantage of the ARMA approximation (i.e., for a given choice of $r=p$, if we set $q<p$ then we are forcing some MA coefficients to be 0, whereas they could help with a better fit if we leave them unrestricted.)

Finally, as for the projection of ideal prices on other firms' mistakes, noting that $x^v_{j,k,t}$ is already a stationary process that is transitory due to the i.i.d. nature of signal noises, we can continue to approximate it with an MA(T_v) process as in Equation (J.50). Thus, the state space representation of the model (under an AR(I)MA approximation of the monetary block) is given by:

$$U^t_{j,k}\equiv\begin{bmatrix} \Xi_{j,t} \\ V^t_{j,-k}\end{bmatrix}=\underbrace{\begin{bmatrix} A_{j,u} & 0_{(r+1)\times T_v} \\ 0_{T_v\times(r+1)} & A_v\end{bmatrix}}_{\equiv A_j}U^{t-1}_{j,k}+\underbrace{\begin{bmatrix} Q_u & 0_{(r+1)\times 1} \\ 0_{T_v\times 1} & Q_v\end{bmatrix}}_{\equiv Q}\begin{bmatrix} u_t \\ v_{j,-k,t}\end{bmatrix} \tag{J.70}$$

with the following implied approximation for $p^*_{j,k,t}$:

$$p^*_{j,k,t}=H'_jU^t_{j,k} \text{ where } H'_j=(\theta_{j,0},\theta_{j,1},\dots,\theta_{j,r},\alpha_j\psi_{K_j,v,0},\dots,\alpha_j\psi_{K_j,v,T_v}) \tag{J.71}$$

Thus, we can modify Algorithm 1 to solve the model with AR(I)MA approximation of the state space representation as follows:

Algorithm 2 (Solving the Model with ARMA Approximation)**.** For a given $K \in \text{Supp}(\mathcal{K})$:

1. Start with a guess for $(\Delta\psi_{K,u,\tau})_{\tau=0}^{T_u}, (\psi_{K,v,\tau})_{\tau=0}^{T_v}$ (at iteration 0, set them equal to their values under rational expectations with full information: $\Delta\psi_{K,u,\tau} = \rho^\tau$ and $\psi_{K,v,\tau} = 0$). Compared to Algorithm 1, choose T_u to be large.[71]
2. Approximate $(\Delta\psi_{K,u,\tau})_{\tau=0}^{T_u}$ with an ARMA(r,r) process (here I use the `'z-Tran'` package of Tan and Wu (2023) to do so, which I describe in more detail below). Choose r to be large enough that the approximation is accurate (I find $r = 5$ is more than sufficient and use this value).
3. Given the ARMA approximation from the previous step, form the state space representation for $x_{j,t}^u$ and $x_{j,k,t}^v$ from Equations (J.50) and (J.68). Then, form matrices A_j, Q, and H_j using Equations (J.70) and (J.71). Solve the rational inattention problem of the firm in Equation (J.19) using the method in Afrouzi and Yang (2019).
4. Using the solution to the dynamic rational inattention problem, find the stationary pair $(\Sigma_{j,-1}, \Sigma_{j,0})$ and the implied Y_j, Λ_j, and $\sigma_{e,j}^2$, construct the time-invariant IRFs of a firm's optimal price with respect to shocks in Equations (J.27) and (J.28).
5. Use fixed point Equations (J.38) and (J.39) to derive the implied $(\psi_{K,u,\tau})_{\tau=0}^{T_u}, (\psi_{K,v,\tau})_{\tau=0}^{T_v}$.
6. Update the guess for $(\Delta\psi_{K,u,\tau})_{\tau=0}^{T_u}, (\psi_{K,v,\tau})_{\tau=0}^{T_v}$ using the implied values from Step 5 and repeat the process in Steps 1 to 6 until convergence.
7. Once the fixed point is found for all $K \in \text{Supp}(\mathcal{K})$, construct the aggregate IRFs of prices and output using Equations (J.41) and (J.42).

Implementation of Algorithm 2 in the Replication Package. The file `solve_model_arima.m` in `./matlab/codes/` folder solves the model for a given set of parameters using Algorithm 2 and is called by `./codes/matlab/solve_model_arima.m`, which takes an array of values of $K \in \text{Supp}(\mathcal{K})$ as well as other parameters (e.g., multiple values of ω) and dispatches multiple instances of the first file for parallel computation. `./codes/matlab/solve_model_arima.m` has a similar structure to `./codes/matlab/solve_model_int_ma.m`, with the exception that when constructing the state space, instead of the integrated MA state space, it calls another internal function, `arima_approx.m`, that performs the ARMA(r,r) approximation described above, and constructs the ARIMA representation in Equation (J.68). It then returns the matrices $A_{j,u}, H_{j,u}$ and Q_u. The matrices are then augmented according to Equation (J.71) after which the rest of the algorithm proceeds as in Algorithm 2 to solve the rational inattention problem using the `DRIPs.m` package of Afrouzi and Yang (2019).

As for the ARMA(r,r) approximation, I use the `'z-Tran'` package of Tan and Wu (2023) that is based on theoretical results from Han et al. (2022), which shows that ARMA processes are dense among stationary processes of the type that we seek to approximate and offers an approximation result. Within `arima.m`, `'z-Tran'` is called through its `'eval'` and `'varma.fit'` functions, which take as input a

[71] For this algorithm, I choose $T_u = 100$ quarters, thus conjecturing that any effects of monetary policy shocks on inflation and output should die out and converge to zero within 25 years. I confirm this is true in the solution of the calibrated model: the IRFs of inflation and output are zero after around 12 to 15 quarters (3 to 4 years), as shown in Figure A.5.

vector of coefficients of the lag polynomials (IRFs) of the process and returns the coefficients of the approximated ARMA process.

K Analytical Decomposition of Strategic Inattention vs. Real Rigidities in the Static Model

The analytical framework of the static model with endogenous capacity in Section 2.4 provides an appropriate framework to discuss the interaction of the strategic complementarity channel with the strategic inattention channel. In particular, recall from Equation (5) that in a symmetric equilibrium, the average price across oligopolies with K competitors, p_K, is given by $\delta_K q$ where

$$\delta_K = \frac{(1-\alpha_K)\lambda_K}{1-\alpha_K\lambda_K} \tag{K.1}$$

where we have now indexed δ, α and λ with K to emphasize that in the micro-founded model, all three of these objects vary with the number of firms in the oligopoly: α_K depends on K through the demand structure, as derived in Equation (24), while λ_K depends on K both through the micro-foundations of the curvature of the profit function B_K, as in Equation (28), as well as through the equilibrium forces as shown in Proposition 4.

Now, defining output of sectors with K competitors as the difference between nominal demand and their average price, $y_K = q - p_K$, it follows that the response of output to the monetary shock q is given by $1-\delta_K$. Thus,

$$\partial y_K/\partial q = 1-\delta_K = \frac{1-\lambda_K}{1-\alpha_K\lambda_K} \tag{K.2}$$

Therefore, the question of how monetary non-neutrality changes with K maps to how $1-\delta_K$ varies with K. In particular, differentiating Equation (K.2) with respect to K formalizes the role of the strategic complementarity and the strategic inattention channels in determining the response of output to monetary policy:

$$\partial_K(\partial y_K/\partial q) = \underbrace{\frac{(1-\lambda_K)\lambda_K}{(1-\alpha_K\lambda_K)^2}\partial_K\alpha_K}_{\text{Channel A: Strategic Complementarity}} - \underbrace{\frac{1-\alpha_K}{(1-\alpha_K\lambda_K)^2}\partial_K\lambda_K}_{\text{Channel B: Strategic Inattention}} \tag{K.3}$$

This decomposition shows that (a) fixing λ_K, a higher α_K increases monetary non-neutrality (Channel A), and (b) fixing α_K, a higher λ_K decreases monetary non-neutrality (Channel B). Thus, the question of how monetary non-neutrality is affected by K boils down to how α_K and λ_K vary with K.

Elasticities of α_K and λ_K with respect to K. The question of how α_K moves with K is related to how demand elasticities vary with firms' market shares, as discussed in Section 4.2 and Equation (24) (See also Appendix F.2 for how α_K and B_K depend on the curvature of a general profit function, or Appendix G for the form of α_K under a Kimball aggregator).

How λ_K moves with K, however, is more complex because it depends on the endogenous attention strategy of firms. As discussed in Section 2.4 and in particular Equations (9) and (10), in a symmetric equilibrium with strictly positive capacity,

$$\lambda_K = \lambda_K(\omega/B_K, V_K^*) = 1 - \frac{\omega}{B_K V_K^*} \tag{K.4}$$

where $V_K^* = V^*(\omega/B_K, \alpha_K, K)$ itself depends on parameters K, B_K and λ_K. Thus,

$$\partial_K \lambda_K = (1-\lambda_K)(\partial \ln(B_K) + \partial_K \ln(V_K^*)) \tag{K.5}$$

where the first term is the direct effect of how the curvature of firms' profit function changes with K. As shown in Equation (28) and derived in Appendix F.2, for a general demand structure, this term depends on the demand elasticity ε_D^K and on the pass-through $1-\alpha_K$: $B_K = \frac{\varepsilon_D^K}{1-\alpha_K}$ where ε_D^K is the demand elasticity of a firm with K competitors. Note that the curvature of the profit function, B_K, increases with the demand elasticity and the degree of strategic complementarity itself. Thus, independent of its direct effect on firms' prices, strategic complementarity also has an impact on firms' strategic inattention through the curvature of their profit functions:

$$\partial_K \ln(B_K) = \underbrace{\partial_K \ln(\varepsilon_D^K)}_{\text{change in elasticity w.r.t. } K} + \underbrace{\frac{1}{1-\alpha_K}\partial_K \alpha_K}_{\text{change in pass-through w.r.t. } K} \tag{K.6}$$

As for the second term in Equation (K.5), it captures the equilibrium effects of B_K, α_K, and K on the prior variance of firms' ideal prices, V^*, which is characterized in Appendices C.5 and C.6.

Now, plugging Equations (K.5) and (K.6) into Equation (K.3), we get at the following decomposition for the total effect of K on monetary non-neutrality through Channels A and B:

$$\partial_K(\partial y_K/\partial q) = \underbrace{\frac{(1-\lambda_K)\lambda_K}{(1-\alpha_K\lambda_K)^2}\partial_K \alpha_K}_{\text{Channel A}} \underbrace{- \overbrace{\frac{1-\lambda_K}{(1-\alpha_K\lambda_K)^2}\partial_K \alpha_K}^{\text{pass-through on } B_K} - \frac{(1-\alpha_K)(1-\lambda_K)}{(1-\alpha_K\lambda_K)^2}\Big[\overbrace{\partial_K \ln(\varepsilon_D^K)}^{\text{elasticity on } B_K} + \overbrace{\partial_K \ln(V_K^*)}^{V_K^* \text{ on } \lambda_K}\Big]}_{\text{Channel B}} \tag{K.7}$$

To unpack this decomposition, K affects monetary non-neutrality through three objects. First, it affects results through how α_K changes with K, which shows up in both Channels A (the real rigidity channel) and B (by affecting strategic inattention of firms through the curvature of profit function). Second, K also affects the results through how it changes the elasticity ε_D^K, which shows up in Channel B by affecting the curvature of firms' profit function. Finally, K also affects monetary non-neutrality by changing V_K^*, which is an equilibrium object and itself depends on K, B_K and α_K.

First-Order Effects of K on Monetary Non-Neutrality. To investigate Equation (K.7) analytically, let us do a Taylor expansion of Equation (K.7) around $\omega/B = 0$, as discussed in Section 2.4 and derived in Appendix C.8. In particular, to simplify the expressions above, let us consider the first order effects of ω by using the results from Appendix C.8 and plugging $B_K = \frac{\varepsilon_D^K}{1-\alpha_K}$, in which case:

$$\partial_K(\partial y_K/\partial q) = \underbrace{\frac{\omega}{\varepsilon_D^K(1-\alpha_K)}\partial_K\alpha_K}_{\text{Channel A (first-order effects of }\omega)} \underbrace{-\overbrace{\frac{\omega}{\varepsilon_D^K(1-\alpha_K)}\partial_K\alpha_K}^{\text{pass-through on }B_K} - \frac{\omega}{\varepsilon_D^K}\overbrace{\partial_K \ln(\varepsilon_D^K)}^{\text{elasticity on }B_K}}_{\text{Channel B (first-order effects of }\omega)} + \mathcal{O}(\|\frac{\omega}{B_K}\|^2) \quad \text{(K.8)}$$

$$= \underbrace{-\frac{\omega}{\varepsilon_D^K}\partial_K \ln(\varepsilon_D^K)}_{\text{total first-order effect}} + \mathcal{O}(\|\frac{\omega}{B_K}\|^2) \quad \text{(K.9)}$$

where $\mathcal{O}(\|\omega\|^2)$ contains the second order terms, including the effect of K on V_K^*. It is important to note that the two effects of α_K cancel out up to first order: fixing capacity, a higher α_K increases monetary non-neutrality through the real rigidity channel (Channel A). However, with endogenous capacity, this effect is offset up to first order as a higher α_K also increases the curvature of firms' profit functions and motivates firms to pay more attention to the fundamental shocks.

Thus, the total first-order effect of how K affects the response of output to the monetary shocks depends *only* on how the demand elasticity changes with K. Everything else, including $\partial\alpha_K/\partial K$, is of higher order in ω, which I discuss more below. As for how demand elasticity changes with K, the theory predicts that firms with more competitors have higher elasticities, and lower markups. For instance, Atkeson and Burstein (2008)'s model implies that demand elasticity should decrease with market share and increase with K (recall that market share in the symmetric equilibrium is $1/K$). There is also empirical evidence for this prediction (for recent evidence, see, e.g., Burstein, Carvalho, & Grassi, 2020; Burya & Mishra, 2023). Therefore, given the positive sign of $\partial_K \ln(\varepsilon_D^K)$, we arrive at the conclusion that, for a general demand structure, strategic inattention channel dominates the real rigidity channel up to first order in ω, and thus monetary non-neutrality decreases with K up to first order in ω.

This result hinges on the fact that while ∂_K/α_K has first-order effects on both Channels A and B, these first-order effects are perfectly symmetric and cancel out. This, of course, raises two questions: (1) why is the effect of α_K on monetary non-neutrality second-order? and (2) what would dampen the effect of Channel B or even make Channel A dominate?

The answer to the first question is that firms' endogenous capacity is maximally sensitive to changes in α_K. In other words, with a larger α_K they increase their information processing capacity so much that it offsets its real rigidity effects. Thus, to break this result, we would need firms to be less responsive in their choice of capacity to changes in the curvature of their profit functions. Moreover, since the benefit of choosing a higher λ_K is derived under a general demand structure and only assumes differentiability of demand, the answer to the second question must rely on the structure of the cost of attention. In particular, the extent to which λ_K responds to the higher curvature introduced by α_K is regulated by the assumptions on the curvature of the cost of attention. The baseline assumption of the rational inattention literature that the cost of attention is linear is Shannon's mutual information (i.e., linear in κ_K), which is the cost function in this paper as well. This is therefore the key assumption that delivers the strong response of λ_K to an increase in α_K. Thus, to eliminate the dominance or generally dampen the strength of Channel

B, one needs to introduce convexity to the cost of information function to dampen the responsiveness of firms to changes in the benefits of information acquisition.

Theoretical and empirical research on the convexity of information costs in rational inattention models is sparse. Nonetheless, the little evidence that we have suggests that the linear cost seems to be a better fit to the data than convex costs. For instance, (Afrouzi & Yang, 2021) show that in a dynamic problem, higher convexity of the cost function in Shannon's mutual information translates to more "smoothing" of Kalman gains over time. However, in the New Zealand survey, learning is lumpy as firms do not acquire information until they need it, indicating that the linear cost fits better with the evidence.

L Dynamic Model with Atkeson and Burstein (2008) Preferences

In this section, I solve the dynamic model where strategic complementarity decreases with K. In particular, I adopt the following preferences based on Atkeson and Burstein (2008):

$$C_t \equiv \left(J^{-1}\sum_{j\in J} C_{j,t}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \quad C_{j,t} \equiv \left(K_j^{-1}\sum_{k\in K_j} C_{j,k,t}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \tag{L.1}$$

where $\sigma < \eta$ so that goods within oligopolies are closer substitutes than goods across oligopolies. The preferences in the benchmark model correspond to the special case where $\sigma \downarrow 1$. It follows that the demand elasticity of firm j,k is given by

$$\varepsilon_{j,k,t}^D = \sigma m_{j,k,t} + \eta(1-m_{j,k,t}) \tag{L.2}$$

where $m_{j,k,t}$ is the market share of firm j,k in period t within oligopoly j as in Equation (22). To obtain the expression for strategic complementarity in this case, we can again differentiate the best response of the firm, similar to Equation (23), to get

$$\alpha_{j,k,t}^{\gamma=0} = \frac{(1-\eta^{-1})m_{j,k,t}}{\frac{\sigma-1}{\eta-\sigma}(1+m_{j,k,t}+\frac{\sigma}{\eta}\frac{m_{j,k,t}^2}{1-m_{j,k,t}})+1} \tag{L.3}$$

We can then proceed to derive the equivalent of the approximate rational inattention problem of the firms (Equation (26)) in this setting as:

$$\max_{\{\kappa_{j,k,t},S_{j,k,t},p_{j,k,t}(S_{j,k}^t)\}_{t\geq 0}} -\texttt{rs}_j \mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t \Big(\underbrace{\frac{1}{2}B_j(p_{j,k,t}(S_{j,k}^t)-p_{j,k,t}^*)^2}_{\text{loss from mispricing}} + \underbrace{\omega\kappa_{j,k,t}}_{\text{cost of capacity}} |S_{j,k}^{-1}\Big)\right] \tag{L.4}$$

$$\begin{aligned} s.t. \quad & p_{j,k,t}^* \equiv (1-\alpha_j)q_t + \alpha_j p_{j,-k,t}(S_{j,-k,t}) \\ & \mathcal{I}\Big(S_{j,k,t},(q_\tau,p_{l,m,\tau}(S_{l,m}^\tau))_{0\leq\tau\leq t}^{(l,m)\neq(j,k)}\Big) \leq \kappa_{j,k,t}, \quad S_{j,k}^t = S_{j,k}^{t-1}\cup S_{j,k,t}, \quad S_{j,k}^{-1} \text{ given.} \end{aligned}$$

where

$$\alpha_j = \frac{(1-\eta^{-1})K_j^{-1}}{\frac{\sigma-1}{\eta-\sigma}(1+K_j^{-1}+\frac{\sigma}{\eta}\frac{K_j^{-2}}{1-K_j^{-1}})+1}, \quad B_j = \frac{\varepsilon_j^D}{1-\alpha_j} = \frac{(\eta-\sigma)(\eta-1)(1-K_j^{-1})K_j^{-1}}{\sigma-1+(\eta-\sigma)(1-K_j^{-1})} + \eta - (\eta-\sigma)K_j^{-1} \tag{L.5}$$

Having characterized this problem, we can then use the same solution method to solve the model and obtain the equilibrium processes for the prices of all sectors. As before we let $p_{k,t} = \mathbb{E}^j[p_{j,t}|K_j = k]$ denote the average log-price implied by this solution across sectors with k competitors. The only difference is how we would compute the output of sectors since the elasticity of substitution across sectors (σ) is now potentially larger than 1. To do this, we only need to use the total demand function of sectors which for any j is given by:

$$Y_{j,t} = Y_t(P_{j,t}/P_t)^{-\sigma} \Rightarrow y_{j,t} = y_t - \sigma(p_{j,t} - p_t) \tag{L.6}$$

where Y_t is the aggregate output, P_t is the aggregate price index and small letters denote log deviations from the steady state. Using the fact that $Q_t = P_t Y_t$ we can then calculate the average output of sectors with k competitors as:

$$y_{k,t} \equiv \mathbb{E}^j[y_{j,t}|K_j = k] = q_t + (\sigma - 1)p_t - \sigma p_{k,t} = \sigma(q_t - p_{k,t}) - (\sigma - 1)y_t \tag{L.7}$$

To calibrate the values of η and σ in the model, I use the following relationship between markups and the number of competitors in the model:

$$\mu_K = \frac{\eta - (\eta - \sigma)K^{-1}}{\eta - 1 - (\eta - \sigma)K^{-1}} \Rightarrow 1/(\mu_K - 1) = \sigma - 1 + (\eta - \sigma)(1 - K^{-1}) \tag{L.8}$$

I then use the survey question that asks firms about their average markups (see the discussion of Table I.1) as well as the survey question that asks about the number of their competitors to generate the variables $1/(\mu_i - 1)$ and $1 - K_i^{-1}$ where i denotes a firm in the survey. I then regress $1/(\mu_i - 1)$ on $1 - K_i^{-1}$. According to the relationship derived from the model, the constant of this regression in Column (2)—which is 1.74—should give us $\sigma - 1$ and the coefficient on $1 - K_i^{-1}$—which is 3.4—should give us $(\eta - \sigma)$ as shown in Table L.1. The resulting values are $\sigma = 2.74$ and $\eta = 6.14$. Moreover, to ensure that strategic complementarity is decreasing in K, I assume that $\gamma = 0$. Given these values of σ, η, and γ, the strategic complementarity in this model *decreases* with K from 0.22 at $K = 2$ to 0 as $K \to \infty$.[72] As for the other parameters, I calibrate them to the same moments in Table 3. In particular, I choose ω in each model to match the coefficient in Table A.2.

Table L.2 shows the results from this exercise for output and inflation responses. Column (2) shows monetary non-neutrality decreases with K even though α_K decreases with K. For instance, the output response is 2.06 times larger in duopolies relative to the monopolistic competition benchmark and this amplification factor declines as K increases. Consistently, Column (6) shows that inflation is more responsive to monetary shocks as K increases. This is consistent with the analytical decomposition in Equation (34), which showed that monetary non-neutrality should decrease with K, independent of the sign of $\partial_K \alpha_K$, as long as demand elasticities are decreasing in K, which is the case in this model.

Moreover, even though decreasing α_K dampens the response of information processing capacity by reducing the curvature of firms' profit functions, Figure L.1 shows that it is still the case that firms with

[72] As discussed in Section 4.2, with $\gamma = 0$, the model cannot match the levels of strategic complementarity documented in the survey data. However, since our goal is to compare output responses in different sectors with different values of K, the key feature of interest is to parameterize the model such that strategic complementarity decreases with K.

Table L.1: Calibration of η and σ

	(1) $1/(\mu-1)$		(2) $1/(\mu-1)$	
$1-K^{-1}$	2.629	(0.337)	3.405	(0.335)
Manufacturing			-1.046	(0.183)
Professional and Financial Services			-2.315	(0.181)
Trade			-0.599	(0.189)
Other			0.578	(1.220)
Constant	1.126	(0.303)	1.746	(0.309)
Observations	3152		3152	

Standard errors in parentheses

Notes: Column (1) of the table shows the results of the regression of $1/(\mu_i - 1)$ on $1-K_i^{-1}$ in the first wave of the survey from Coibion et al. (2018). Column (2) reports the result of the same regression while controlling for industry fixed effects shown in the table. The constant of the regression corresponds to $\sigma - 1$ in the model while the coefficient on $1 - K_i^{-1}$ corresponds $(\eta - \sigma)$.

larger K produce more capacity and allocate more of it towards aggregates. However, the slope of the increase is much smaller than the benchmark model. Again, the analytical decomposition of the response of capacity in Equation (33) sheds light on this result. The curvature of the profit function is affected by both the sign of $\partial_K \alpha_K$ and the sign of $\partial_K \ln(\varepsilon_D^K)$. Even though a negative $\partial_K \alpha_K$ reduces capacity with K, this effect is dominated by the increase in the curvature of the profit function due to a positive $\partial_K \ln(\varepsilon_D^K)$.

Moreover, Table L.3 shows the decomposition of the change in monetary non-neutrality to the strategic inattention and real rigidity channels as derived in Equation (31). This table shows that (1) with decreasing strategic complementarities, both channels move in the same direction and reduce monetary non-neutrality and (2) the share of the strategic inattention channel is smaller because the decreasing strategic complementarities dampen the curvature of the profit function and reduce the response of capacity to monetary shocks. This also can be seen analytically in the static model and in particular in the discussion of Equation (34).

It is also worth pointing out why the amplification factor of the benchmark model relative to the monopolistic competition model is so small in this exercise. This happens because due to the assumption of $\gamma = 0$, which is made to generate the decreasing strategic complementarities in the model, the level of strategic complementarities are small on average across sectors (average α in the benchmark and monopolistic competition model is 0.087 as opposed to 0.817 in the calibrated model matched to survey data). Because the model misses this moment in this calibration, comparing the benchmark and monopolistic competition models is not very informative. Instead, the main value of this exercise is its predictions for the amplification factors across sectors with different K. Nonetheless, I have included this comparison for consistency and completeness.

Finally, to check the robustness of these results to the calibration of η and σ, I redo the analysis while keeping η at its original calibration of 12, and fixing $\sigma = 6$, which is a common calibration of this parameter in monetary models. In this case, strategic complementarity declines from 0.186 when $K=2$ to 0 as $K\to\infty$. The results are presented in Tables L.4 and L.5 and Figure L.2 and are qualitatively similar to the results discussed above. Quantitatively, the amplification factors are larger and more dispersed because with higher elasticities of substitution, small differences in responsiveness of prices translate to stronger reallocation of demand *across* sectors.

Table L.2: Robustness — Atkeson and Burstein (2008) Preferences

Model		*Output*				*Inflation*			
		Variance		*Persistence*		*Variance*		*Persistence*	
		var(Y) $^{\times 10^4}$	*amp. factor*	*half-life* ^{qtrs}	*amp. factor*	*var(π)* $^{\times 10^4}$	*damp. factor*	*half-life* ^{qtrs}	*amp. factor*
		(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Monopolistic Competition		1.40	1.00	3.45	1.00	1.60	1.00	3.88	1.00
Benchmark	$K \sim \hat{\mathcal{K}}$	1.42	1.01	3.47	1.01	1.59	1.00	3.89	1.00
2-Competitors	$K=2$	2.88	2.06	3.86	1.12	1.52	0.95	4.06	1.05
4-Competitors	$K=4$	1.69	1.21	3.55	1.03	1.58	0.99	3.93	1.01
8-Competitors	$K=8$	1.25	0.89	3.38	0.98	1.60	1.00	3.87	1.00
16-Competitors	$K=16$	1.07	0.76	3.30	0.96	1.62	1.01	3.84	0.99
32-Competitors	$K=32$	0.98	0.70	3.25	0.94	1.62	1.02	3.83	0.99
∞-Competitors	$K \to \infty$	0.91	0.65	3.21	0.93	1.63	1.02	3.81	0.98

Notes: the table shows robustness statistics for output and inflation responses across models with different number of competitors at the micro-level and presents results for an alternative calibration of the model with constant returns to scale and Atkeson and Burstein (2008) preferences where strategic complementarities decrease with K. $var(.)$ denotes the variance of output/inflation conditional on monetary shocks. *Half-life* denotes the length of the time that it takes for inflation/output to live half of its cumulative response in quarters. *Damp. factor* (*amp. factor*) denotes the factor by which the relevant statistic is smaller (larger) in the corresponding model relative to the model with monopolistic competition.

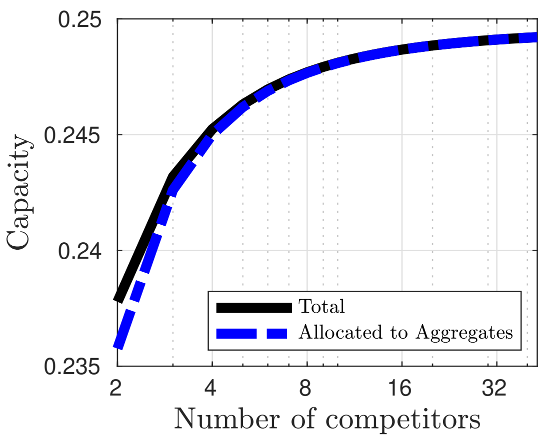


Figure L.1: Information Capacity for Different K.

Notes: the figure shows the produced information processing capacity of a firm as a function of the number of competitors within its sector in the model with Atkeson and Burstein (2008) with low elasticities of substitution. Firms with more competitors acquire more information and allocate more of it toward aggregates.

Table L.3: Decomposition: Strategic Inattention vs. Real Rigidities

	Percentage change in variance of	
	output	*inflation*
	(1)	*(2)*
Total Change (percent)	43.8	-6.8
Due to Str. Inattention (ppt)	5.4	-1.1
Due to Real Rigidities (ppt)	38.5	-5.7

Notes: The table shows the decomposition of the effects of the strategic inattention and real rigidity channels for the change in volatility of output (monetary non-neutrality) and inflation conditional on monetary shocks, as derived in Equation (31), under Atkeson and Burstein (2008) preferences with low elasticities of substitution.

Table L.4: Robustness — Atkeson and Burstein (2008) Preferences with High Elasticities of Substitution

		Output				*Inflation*			
		Variance		*Persistence*		*Variance*		*Persistence*	
Model		*var(Y)* $^{\times 10^4}$	*amp. factor*	*half-life* ^{qtrs}	*amp. factor*	*var(π)* $^{\times 10^4}$	*damp. factor*	*half-life* ^{qtrs}	*amp. factor*
		(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Monopolistic Competition		1.38	1.00	3.45	1.00	1.60	1.00	3.88	1.00
Benchmark	$K \sim \hat{\mathcal{K}}$	1.40	1.02	3.47	1.01	1.59	1.00	3.89	1.00
2-Competitors	$K=2$	4.59	3.32	4.05	1.17	1.53	0.96	4.03	1.04
4-Competitors	$K=4$	1.94	1.41	3.61	1.05	1.58	0.99	3.92	1.01
8-Competitors	$K=8$	1.10	0.80	3.29	0.95	1.60	1.00	3.86	0.99
16-Competitors	$K=16$	0.78	0.57	3.08	0.89	1.61	1.01	3.84	0.99
32-Competitors	$K=32$	0.65	0.47	2.95	0.86	1.62	1.01	3.83	0.99
∞-Competitors	$K \to \infty$	0.54	0.39	2.81	0.81	1.63	1.02	3.82	0.98

Notes: The table shows robustness statistics for output and inflation responses across models with different numbers of competitors at the micro-level and presents results for an alternative calibration of the model with constant returns to scale and Atkeson and Burstein (2008) preferences where strategic complementarities decrease with K. $var(.)$ denotes the variance of output/inflation conditional on monetary shocks in the model. *Half-life* denotes the length of the time that it takes for inflation/output to live half of its cumulative response in quarters. *Damp. factor* (*amp. factor*) denotes the factor by which the relevant statistic is smaller (larger) in the corresponding model relative to the model with monopolistic competition.

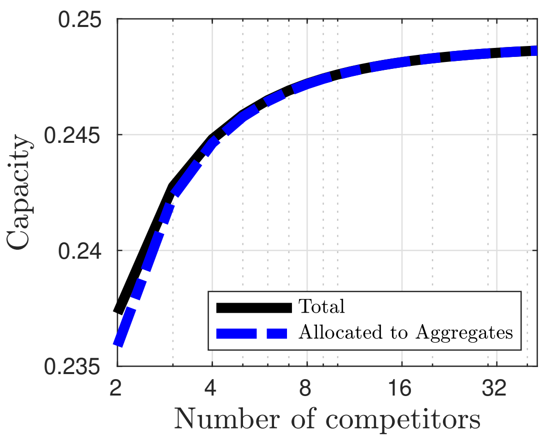


Figure L.2: Information Capacity for Different K.

Notes: the figure shows the produced information processing capacity of a firm as a function of the number of competitors within its sector in the model with Atkeson and Burstein (2008) with high elasticities of substitution. Firms with more competitors acquire more information and allocate more of it toward aggregates.

Table L.5: Decomposition: Strategic Inattention vs. Real Rigidities

	Percentage change in variance of	
	output	*inflation*
	(1)	*(2)*
Total Change (percent)	38.2	-5.9
Due to Str. Inattention (ppt)	6.6	-1.2
Due to Real Rigidities (ppt)	31.6	-4.7

Notes: The table shows the decomposition of the effects of the strategic inattention and real rigidity channels for the change in volatility of output (monetary non-neutrality) and inflation conditional on monetary shocks, as derived in Equation (31), under Atkeson and Burstein (2008) preferences with high elasticities of substitution.

M Additional Robustness Exercises

M.1. Heterogeneity within Sector Market Shares

In the approximate problem that I considered in Section 4.3, all firms have the same market share in the steady state. One question is how heterogeneity in market shares affects strategic inattention. Is it the case that with asymmetries in market shares, larger firms ignore the mistakes of smaller firms, which would potentially dampen strategic inattention? To address this question, I present a simple case with CES preferences and show that the strategic complementarity of any given firm is equal to their market share in the steady-state. This means that firms with higher market shares have higher strategic complementarities and are more likely to pay attention to other firms' mistakes rather than aggregate shocks. Thus, heterogeneity in market share is expected to amplify the effects of strategic inattention.

To see this, consider the household's demand with CES aggregator from Equation 16 with the following modification:

$$C_t = \prod_{j\in J}\left[\left(\sum_{k\in K_j}\bar{m}_{j,k}^{\frac{1}{\eta}} C_{j,k,t}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\right]^{J^{-1}} \tag{M.1}$$

where now $\bar{m}_{j,k}$ captures the taste of the consumer for the product of firm k in industry j. Moreover, $\forall j$ we normalize $\sum_k \bar{m}_{j,k} = 1$ so that that these tastes are relative. It is straight forward to show that $\bar{m}_{j,k}$ shows up as a demand shifter in firm j,k' demand

$$C_{j,k,t} = P_t C_t \frac{\bar{m}_{j,k} P_{j,k,t}^{-\eta}}{\sum_l \bar{m}_{j,l} P_{j,l,t}^{1-\eta}} \tag{M.2}$$

On the firm side, this implies that the elasticity of demand for firm j,k at time t is given by

$$\varepsilon_{j,k,t} = \eta - (\eta-1)\frac{\bar{m}_{j,k} P_{j,k,t}^{1-\eta}}{\sum_l \bar{m}_{j,l} P_{j,l,t}^{1-\eta}} \tag{M.3}$$

On the firm side, assume constant returns to scale in production ($\gamma = 0$) and that there is a subsidy for every firm such that it sets their steady state price equal to the aggregate marginal cost given their optimal markup (so that there is no price dispersion in the steady state). Then the approximate problem of the firm, as in Equation 26, is given by

$$\max_{\{\kappa_{j,k,t}, S_{j,k,t}, p_{j,k,t}(S_{j,k}^t)\}_{t\geq 0}} -\mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t \left(\underbrace{\eta(p_{j,k,t}(S_{j,k}^t) - p_{j,k,t}^*)^2}_{\text{loss from mispricing}} + \underbrace{\omega \kappa_{j,k,t}}_{\text{cost of capacity}} \, | S_{j,k}^{-1}\right)\right] \tag{M.4}$$

$$\begin{aligned} s.t. \quad & p_{j,k,t}^* \equiv (1-\alpha_{j,k}) q_t - \alpha_{j,k} p_{j,-k,t}(S_{j,-k,t}) \\ & \mathcal{I}\left(S_{j,k,t}, (q_\tau, p_{l,m,\tau}(S_{l,m}^\tau))_{0\leq\tau\leq t}^{(l,m)\neq(j,k)}\right) \leq \kappa_{j,k,t} \\ & S_{j,k}^t = S_{j,k}^{t-1} \cup S_{j,k,t}, \quad S_{j,k}^{-1} \text{ given.} \end{aligned} \tag{M.5}$$

where we have already imposed that in the case of $\gamma = 0$, the curvature of the profit function is uniquely determined by the elasticity of substitution ($B_j = \eta$). The only major difference to this problem is that now, with heterogeneity in market shares, there is also heterogeneity in the degree of strategic complementarity within industries. In fact, in this case, the degree of strategic complementarity for every firm is proportional to their steady-state market share:

$$\alpha_{j,k} = (1-\eta^{-1})\bar{m}_{j,k} \tag{M.6}$$

Note that, here, $\bar{m}_{j,k}$ is simply the market share of firm k in industry j in the steady-state, and we can study the impact of heterogeneity in market shares on the attention allocation of firms. Finally, to make this case even simpler, assume that $\eta \to \infty$.[73] Then, taking a second-order approximation around this steady state, it follows from Equation (23) that the ideal price of firm j,k is given by

$$p^*_{j,k,t} = (1-\bar{m}_{j,k})q_t + \bar{m}_{j,k}\frac{\sum_{l\neq k}\bar{m}_{j,l}p_{j,l,t}}{\sum_{l\neq k}\bar{m}_{j,l}} \tag{M.7}$$

This representation also shows that higher market share leads to higher strategic complementarity and hence magnifies the degree of strategic inattention.

M.2. Lower Persistence of Nominal Demand Growth

While many of the parameter values calibrated to the New Zealand data are also consistent with their calibrations for the U.S., one exception is the persistence of the nominal demand growth, ρ. While the value for this parameter is 0.707 in New Zealand, its value in the US is around a monthly persistence of 0.61 (Midrigan, 2011; Mongey, 2021) (or a quarterly persistence of $0.61^3 = 0.23$). To compare the results for this case, I recalibrate the cost of information acquisition and redo the analysis for monetary non-neutrality for $\rho = 0.23$, as shown in Table M.1a in Appendix A. The main takeaway is that while the amplification factors are slightly smaller than the case for $\rho = 0.707$, the results are fairly robust. For instance, relative to the model with monopolistic competition, aggregate output is 23% percent more volatile under the benchmark calibration for the distribution of competitors—as opposed to 28% with $\rho = 0.707$.

M.3. Alternative Discount Factor

One of the mechanisms in attention allocation within the model is firms' dynamic incentives. Forward-looking firms internalize the long-term benefits of learning about more persistent shocks and adjust their information acquisition accordingly (see, e.g., Afrouzi & Yang, 2019). In the model, this mechanism dampens monetary non-neutrality because the dynamic incentives are very strong with $\beta = 0.96^{0.25}$ and mistakes are more transitory than fundamental shocks. To show the strength of this mechanism in dampening strategic inattention, I recalibrate β and ω by jointly targeting the coefficient in Table 1 in addition to the original moment of calibration from Table A.2 and redo the results for monetary non-neutrality, shown in Table M.1b in at the end of this section. The main takeaway is that the effects of strategic inattention is larger in this calibration. The key intuition for these results is that the calibrated

[73] In this hypothetical example, having $\eta \to \infty$ means that firms' profit functions are infinitely concave and that the benefit of information is arbitrarily large given a fixed ω. Therefore, for a fixed ω firms will acquire almost perfect information. To resolve this, we assume that ω is also proportional to η so that the ratio stays constant as $\eta \to \infty$.

β is smaller in this case, which leads to firms producing less capacity and allocating more of it towards the mistakes of their competitors, both of which amplify the effects of strategic inattention.

M.4. The Role of Firm- and Sector-level Idiosyncratic Shocks

In the benchmark model, I have abstracted away from firm- and sector-level idiosyncratic shocks. One question is how such shocks would interact with firms' strategic inattention motives.

In terms of firm-level shocks, they should have a similar effect as competitors' mistakes: firms need to pay attention to others' mistakes as well as shocks to their competitors' costs. Thus, the conjecture is that such shocks would amplify the incentives of firms with fewer competitors to pay less attention to aggregate shocks and more attention to a weighted average of fundamental cost shocks of their competitors. However, as the number of competitors increases, for any given firm, the average fundamental cost shock of their competitors would become smaller due to the law of large numbers which would reduce the firms' incentive to pay attention to others' cost shocks and pay more attention to the aggregate/common shocks within the oligopoly. Thus, I would expect the presence of firm-level fundamental cost shocks to amplify the differential incentives across firms and lead to potentially larger differences in attention allocation across firms with different numbers of competitors.

On the other hand, industry-wide shocks would have a very similar effect as fundamental shocks in the model, as they would be common to all firms in an oligopoly. In fact, in the problem of firms, q can be interpreted either as an aggregate shock or an industry-wide shock. It is only in the process of aggregation that these two differ. To illustrate the effect of such shocks, I have solved a numerical example of the model with these types of shocks. In this numerical example, keeping all parameters the same as the benchmark calibration and setting the standard deviation of i.i.d. (over time) idiosyncratic shocks to twice the standard deviation of fundamental shocks, I have recalibrated the cost of attention to match the same moment in Table 3. The results are presented in Tables M.2 and M.3 and Figure M.1. The presence of i.i.d. idiosyncratic shocks increases the amplification factors in monetary non-neutrality across sectors with different K. The amplification result stems from the assumption that sector-level shocks are i.i.d. over time. Now that common shocks are more transitory; firms assign a smaller continuation value to attending to common shocks, which, fixing ω, reduces their overall attention to q.

Consequently, I believe the model without idiosyncratic shocks provides a conservative benchmark for the effect of competition on attention as far as these types of cost shocks are concerned.

Table M.1: Robustness — Output and Inflation Across Models

(a) Alternative persistence for the growth of nominal aggregate demand ($\rho=0.23$)

		Output				*Inflation*			
		Variance		*Persistence*		*Variance*		*Persistence*	
Model		*var(Y)* $^{\times 10^4}$	*amp. factor*	*half-life* ^{qtrs}	*amp. factor*	*var(π)* $^{\times 10^5}$	*damp. factor*	*half-life* ^{qtrs}	*amp. factor*
		(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Monopolistic Competition		2.26	1.00	2.97	1.00	2.48	1.00	3.78	1.00
Benchmark	$K \sim \hat{\mathcal{K}}$	2.77	1.23	3.31	1.11	2.15	0.87	4.20	1.11
2-Competitors	$K=2$	3.10	1.37	3.73	1.26	1.94	0.78	4.47	1.18
4-Competitors	$K=4$	2.81	1.25	3.37	1.13	2.12	0.86	4.24	1.12
8-Competitors	$K=8$	2.73	1.21	3.25	1.09	2.19	0.88	4.17	1.10
16-Competitors	$K=16$	2.70	1.20	3.20	1.08	2.22	0.89	4.14	1.10
32-Competitors	$K=32$	2.69	1.19	3.18	1.07	2.23	0.90	4.12	1.09
∞-Competitors	$K\to\infty$	2.67	1.18	3.15	1.06	2.25	0.91	4.11	1.09

(b) Alternative discount rate for information (β=0.60)

		Output				*Inflation*			
		Variance		*Persistence*		*Variance*		*Persistence*	
Model		*var(Y)* $^{\times 10^4}$	*amp. factor*	*half-life* ^{qtrs}	*amp. factor*	*var(π)* $^{\times 10^4}$	*damp. factor*	*half-life* ^{qtrs}	*amp. factor*
		(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Monopolistic Competition		3.49	1.00	3.07	1.00	1.57	1.00	4.43	1.00
Benchmark	$K \sim \hat{\mathcal{K}}$	4.76	1.36	3.41	1.11	1.45	0.92	4.76	1.07
2-Competitors	$K=2$	5.80	1.66	3.87	1.26	1.31	0.83	5.02	1.13
4-Competitors	$K=4$	4.89	1.40	3.48	1.13	1.42	0.91	4.80	1.08
8-Competitors	$K=8$	4.65	1.33	3.34	1.09	1.47	0.94	4.73	1.07
16-Competitors	$K=16$	4.55	1.30	3.28	1.07	1.49	0.95	4.70	1.06
32-Competitors	$K=32$	4.51	1.29	3.26	1.06	1.50	0.96	4.69	1.06
∞-Competitors	$K\to\infty$	4.48	1.28	3.23	1.05	1.51	0.96	4.68	1.06

Notes: the table presents robustness statistics for output and inflation responses across models with different number of competitors at the micro-level. Panel (a) presents results for an alternative calibration of persistence in the growth of nominal demand ($\rho=0.23$). Panel (b) presents results for an alternative calibration of discount rate for information ($\beta=0.60$). $var(.)$ denotes the variance of output/inflation conditional on monetary shocks. *Half-life* denotes the length of the time that it takes for inflation/output to live half of its cumulative response in quarters. *Damp. factor* (*amp. factor*) denotes the factor by which the relevant statistic is smaller (larger) in the corresponding model relative to the model with monopolistic competition.

Table M.2: Robustness — Model with Idiosyncratic Shocks within Sectors

Model		*Output*				*Inflation*			
		Variance		*Persistence*		*Variance*		*Persistence*	
		var(Y) $^{\times 10^4}$	*amp. factor*	*half-life* ^{qtrs}	*amp. factor*	*var(π)* $^{\times 10^4}$	*damp. factor*	*half-life* ^{qtrs}	*amp. factor*
		(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Monopolistic Competition		3.04	1.00	3.39	1.00	1.48	1.00	4.39	1.00
Benchmark	$K \sim \hat{\mathcal{K}}$	3.93	1.29	3.70	1.09	1.38	0.94	4.64	1.06
2-Competitors	$K=2$	4.54	1.49	4.12	1.22	1.29	0.88	4.80	1.09
4-Competitors	$K=4$	4.00	1.32	3.76	1.11	1.37	0.93	4.66	1.06
8-Competitors	$K=8$	3.86	1.27	3.63	1.07	1.40	0.95	4.61	1.05
16-Competitors	$K=16$	3.80	1.25	3.58	1.06	1.41	0.96	4.60	1.05
32-Competitors	$K=32$	3.78	1.24	3.55	1.05	1.42	0.96	4.59	1.05
∞-Competitors	$K \to \infty$	3.76	1.24	3.53	1.04	1.42	0.96	4.58	1.04

Notes: the table shows robustness statistics for output and inflation responses across models with different number of competitors in the model with sector level idiosyncratic shocks discussed in Appendix M.4. $var(.)$ denotes the variance of output/inflation conditional on monetary shocks. *Half-life* denotes the length of the time that it takes for inflation/output to live half of its cumulative response in quarters. *Damp. factor* (*amp. factor*) denotes the factor by which the relevant statistic is smaller (larger) in the corresponding model relative to the model with monopolistic competition.

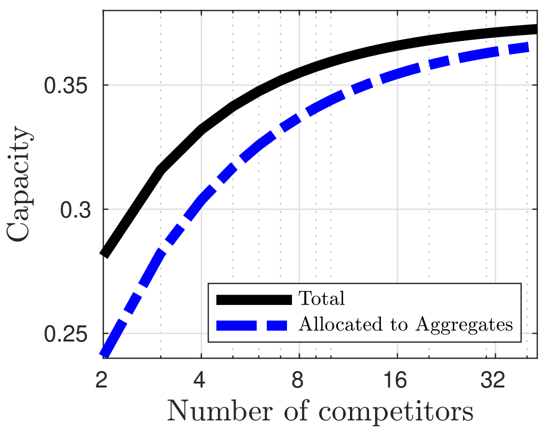


Figure M.1: Information Capacity for Different K.

Notes: the figure shows the produced information processing capacity of a firm as a function of the number of competitors within its sector in the model with sector-level idiosyncratic shocks in Appendix M.4. Firms with more competitors acquire more information and allocate more of it toward aggregates.

Table M.3: Decomposition: Strategic Inattention vs. Real Rigidities

	Percentage change in variance of	
	output	*inflation*
	(1)	*(2)*
Total Change (percent)	18.9	-9.6
Due to Str. Inattention (ppt)	82.0	-19.7
Due to Real Rigidities (ppt)	-63.1	10.1

Notes: The table shows the decomposition of the effects of the strategic inattention and real rigidity channels for the change in volatility of output (monetary non-neutrality) and inflation conditional on monetary shocks, as derived in Equation (31) in the model with sector-level idiosyncratic shocks in Appendix M.4.

References

Afrouzi, H., & Yang, C. (2019). *Dynamic rational inattention and the phillips curve.* (Manuscript)

Afrouzi, H., & Yang, C. (2021). *Selection in information acquisition and monetary non-neutrality.* (Manuscript)

Amiti, M., Itskhoki, O., & Konings, J. (2019). International shocks, variable markups, and domestic prices. *The Review of Economic Studies*, *86*(6), 2356–2402.

Atkeson, A., & Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. *American Economic Review*, *98*(5), 1998–2031.

Bilal, A. G., Engbom, N., Mongey, S., & Violante, G. L. (2019). *Firm and worker dynamics in a frictional labor market.* (Manuscript)

Burstein, A., Carvalho, V. M., & Grassi, B. (2020). *Bottom-up markup fluctuations* (Working Paper No. 27958). National Bureau of Economic Research. doi: 10.3386/w27958

Burya, A., & Mishra, S. (2023). *Variable markups, demand elasticity and pass-through of marginal costs into prices.* (Manuscript)

Coibion, O., Gorodnichenko, Y., & Kumar, S. (2018). How do firms form their expectations? new survey evidence. *American Economic Review*, *108*(9), 2671–2713.

Coibion, O., Gorodnichenko, Y., Kumar, S., & Ryngaert, J. (2021). Do you know that I know that you know...? Higher-order beliefs in survey data. *The Quarterly Journal of Economics*, *136*(3), 1387–1446.

Cover, T. M., & Thomas, J. A. (2012). *Elements of information theory*. John Wiley & Sons.

Hamilton, J. D. (1994). *Time series analysis*. Princeton university press.

Han, Z., Tan, F., & Wu, J. (2022). Analytic policy function iteration. *Journal of Economic Theory*, *200*, 105395.

Maćkowiak, B., Matějka, F., & Wiederholt, M. (2018). Dynamic rational inattention: Analytical results. *Journal of Economic Theory*, *176*, 650 - 692.

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

Maćkowiak, B., & Wiederholt, M. (2009). Optimal sticky prices under rational inattention. *The American Economic Review*, *99*(3), 769–803.

Midrigan, V. (2011). Menu costs, multiproduct firms, and aggregate fluctuations. *Econometrica*, *79*(4), 1139–1180.

Mongey, S. (2021). *Market structure and monetary non-neutrality* (Working Paper No. 29233). National Bureau of Economic Research.

Romer, C. D., & Romer, D. H. (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review*, *94*(4), 1055–1084.

Tan, F., & Wu, J. (2023). *z-tran: Matlab software for solving dynamic incomplete information models.* Available at GitHub: `https://github.com/econdojo/ztran`.

Wang, O., & Werning, I. (2022). Dynamic oligopoly and price stickiness. *American Economic Review*, *112*(8), 2815–2849.