

Strategic Inattention, Inflation Dynamics, and the Non-Neutrality of Money*

Hassan Afrouzi[†]
Columbia University

First Draft: Nov 17, 2016

This Draft: Mar 28, 2020

Abstract

How does competition affect information acquisition of firms and thus the response of inflation and output to monetary policy shocks? This paper addresses these questions in a new dynamic general equilibrium model with both dynamic rational inattention and oligopolistic competition. In the model, rationally inattentive firms acquire information about the endogenous beliefs of their competitors. Moreover, firms with fewer competitors endogenously choose to acquire less information about aggregate shocks – a novel prediction of the model that is supported by empirical evidence from survey data. A quantitative exercise disciplined by firm-level survey data shows that firms’ strategic inattention to aggregate shocks associated with oligopolistic competition increases monetary non-neutrality by up to 77% and amplifies the half-life of output response to monetary shocks by up to 30%. Furthermore, the model matches the relationship between the number of firms’ competitors and their uncertainty about inflation as a non-targeted moment.

JEL classification: E31; E32; E71

Key Words: rational inattention, oligopolistic competition, inflation dynamics, inflation expectations, monetary non-neutrality.

*This paper is a revised version of the first chapter of my PhD dissertation at UT Austin. I am grateful to Olivier Coibion, Saroj Bhattarai and Yuriy Gorodnichenko for their guidance and support, and Ricardo Reis and Raphael Schoenle for thoughtful discussions of earlier drafts of this paper. I would also like to thank V. Bhaskar, Nicholas Buchholz, Mark Dean, Andres Drenik, Matthieu Gomez, Andrew Glover, Guido Lorenzoni, Filip Matejka, Chris Moser, Eugenio Miravete, Emi Nakamura, Jesse Schreger, Christopher Sims, Jón Steinsson, Max Stinchcombe, Mirko Wiederholt, Emily Weisburst, Michael Woodford, Pierre Yared and numerous seminar participants at several different institutions.

[†]Email: hassan.afrouzi@columbia.edu. Address: 1105A International Affairs Building, 420 West 118th Street, New York, NY 10027.

1 Introduction

Since the seminal work of [Friedman \(1968\)](#) and [Phelps \(1967\)](#), macroeconomists have emphasized the importance of expectations for the evolution of prices in the economy. Almost every modern monetary model relates aggregate price changes to price-setters' expectations about aggregate inflation.¹ However, recent literature documents that firms' inflation expectations are subject to information frictions,² and firms with a larger number of competitors have more accurate forecasts about aggregate inflation ([Coibion, Gorodnichenko and Kumar, 2018](#)). These facts are not consistent with our standard models and raise the following two questions: (1) Why do firms with more competitors have more accurate inflation expectations? (2) How does the interaction between the number of competitors and the accuracy of inflation expectations affect the propagation of monetary policy shocks to output and inflation?

In this paper, I develop a new dynamic general equilibrium model with both *rational inattention* and *oligopolistic competition* to study these questions. In the model, the interaction of these two frictions generates an endogenous correlation between the number of firms' competitors and the accuracy of their inflation expectations. I calibrate the model to firm-level survey data and find that the model matches the relationship between the uncertainty of firms about aggregate inflation and the number of their competitors as a *non-targeted* moment. Finally, I find that the interaction between oligopolistic competition and rational inattention has quantitatively significant implications for the propagation of monetary policy shocks to output and inflation such that it increases monetary non-neutrality by up to 77%.

The model combines rational inattention and oligopolistic competition in an otherwise standard dynamic general equilibrium framework. In the model, households form demand for products of oligopolistic firms and monetary policy sets the aggregate nominal demand. On the firms' side, the economy is populated with a large number of oligopolies and there is heterogeneity in the number of competitors within these oligopolies. All firms are rationally inattentive. They hire labor to produce information processing capacity as well as to produce their consumption good. Furthermore, after choosing how much information processing capacity to produce, they decide how to allocate that capacity between learning about aggregate shocks and the endogenous beliefs of their oligopolistic competitors. Given

¹In New Keynesian models inflation is increasing in expected aggregate inflation in the future ([Woodford, 2003b](#)). In models of information rigidity, it is increasing in past expectation of current inflation ([Lucas Jr, 1972](#); [Mankiw and Reis, 2002](#)).

²For instance, [Kumar, Afrouzi, Coibion and Gorodnichenko \(2015\)](#) document that managers in New Zealand make average errors of 2 to 3 percentage points in perceiving current as well as forecasting future inflation. Similarly, [Bryan, Meyer and Parker \(2015\)](#) document that managers in the U.S. also report much higher as well as more dispersed expectations of overall price changes in the economy.

their optimal information structure, firms then choose their prices and meet their demand.

The first contribution of this paper is to characterize the optimal information structure of rationally inattentive firms in a *strategic* environment within a dynamic general equilibrium model. Firms that compete with only a few others optimize over their price relative to the prices of their direct rivals. When information is costly, such firms find themselves facing an endogenous trade-off: how much to track the exogenous shocks versus the direct beliefs of their rivals about those shocks. In particular, when the number of competitors is finite, the average beliefs of firms' competitors exhibit non-fundamental volatility that is due to endogenous shocks to beliefs. These endogenous shocks, which arise from firms' mistakes in perceiving exogenous shocks, are not only costly to the firms who make them, but also for their competitors due to strategic complementarities in pricing. Accordingly, with costly information, oligopolistic firms have incentives to be *strategically inattentive* to aggregate shocks. They find it optimal to pay direct attention to the mistakes of their competitors, even at the expense of substituting attention away from fundamental shocks that affect their own profits.

The second contribution of this paper is to characterize the dynamic consequences of firms' strategic inattention incentives for the evolution of their beliefs. To do so, I calibrate the dynamic oligopolistic model with rational inattention and heterogeneity in the number of competitors within oligopolies to firm-level survey data. In the calibrated model, firms in oligopolies with a larger number of competitors allocate a higher amount of attention to learning the aggregate shocks. Accordingly, their forecasts about aggregate variables are more accurate, on average, and their posteriors about these variables are more certain. This is a unique prediction of the oligopolistic rational inattention model and is supported by the evidence on point forecasts from [Coibion, Gorodnichenko and Kumar \(2018\)](#). I also provide further evidence for this channel by documenting that firms with a larger number of competitors have more certain posteriors about aggregate inflation (i.e. they are more certain about their forecasts given the distribution of their subjective beliefs). The calibrated model matches the decline in the uncertainty of firms about inflation as a function of their number of competitors as non-targeted moments.

In the model, two channels link information acquisition to the number of competitors within an oligopoly. First, the number of competitors changes the sensitivity of firms' profit functions to *all* shocks and affects the optimal level of information processing capacity that firms choose to produce. This sensitivity is related to firms' elasticities of demand which increase with the number of their competitors. Hence, firms with more competitors produce more capacity for processing information. Second, the number of competitors also changes how firms allocate their produced capacity between aggregate shocks and the endogenous beliefs of their competitors about those shocks. Two separate forces interact for this alloca-

tion of attention: (1) in the calibrated model, profits of firms with more competitors are more sensitive to the mistakes of their rivals because they have larger degrees of strategic complementarities in pricing. This implies an increase in firms' strategic incentives with respect to the number of their competitors. (2) In oligopolies with a larger number of competitors, mistakes wash out due to the law of large numbers and firms' beliefs are on average more correlated with aggregate shocks. Smaller mistakes then diminish the firms' incentives in tracking the beliefs of their competitors.

The resultant effect of competition on information acquisition depends on the interaction of these forces. Overall, in the calibrated model, firms in more competitive oligopolies produce more information processing capacity and allocate a larger amount of that capacity towards learning the aggregate shocks. This prediction of the calibrated model matches the empirical evidence that firms with a larger number of competitors are more informed about aggregate variables (Coibion et al., 2018).

The third contribution of this paper is to quantify the implications of these incentives for the propagation of monetary policy shocks to output and inflation. Since firms in less competitive oligopolies acquire less information about the aggregate shocks, the response of their prices to these shocks are smaller and more persistent, both of which amplify monetary non-neutrality. In a set of counterfactual exercises, I find that this effect is quantitatively significant relative to a model with monopolistic competition: it increases the volatility of output due to monetary shocks by up to 77% and increases the half-life of output by up to 30% (1 quarter). Moreover, it lowers the volatility of inflation caused by monetary shocks by up to 13% and increases its half-life by up to 17% (2 months).

These effects on monetary non-neutrality are driven by two opposing forces in the calibrated model. On the one hand, firms with fewer competitors pay less attention to monetary policy shocks due to strategic incentives, which amplifies monetary non-neutrality through larger information frictions. On the other hand, firms in oligopolies with fewer competitors have lower degrees of strategic complementarity, which attenuates the degree of monetary non-neutrality. A decomposition of the net effects of these two forces shows that while both forces are quantitatively significant in the calibrated model, the first force dominates and the resultant effect is such that monetary non-neutrality is amplified when the number of competitors within oligopolies are smaller.

Finally, to further illustrate the implications of strategic incentives for inflation dynamics, I derive a closed-form expression for the model-implied Phillips curve in the special case where firms are fully myopic and show that it relates inflation primarily to price-setters' expectations about their competitors' price changes rather than their expectations about ag-

gregate inflation.³ Therefore, the model also provides an explanation for why, in countries like New Zealand and the U.S., aggregate inflation can remain low and stable even when price-setters' expectations of aggregate inflation are not. Inflation expectations simply play little role in price-setting decisions when rationally inattentive price-setters have strategic motives.

Literature Review. The paper is mainly and closely related to the literature on rational inattention (Sims, 2003, 2006; Matějka and McKay, 2015), and its implications for business cycles, pricing, and monetary non-neutrality (Maćkowiak and Wiederholt, 2009, 2015; Paciello and Wiederholt, 2014; Matějka, 2015; Pasten and Schoenle, 2016; Stevens, 2020; Yang, 2019).⁴ More broadly, the paper is also related to the literature on the effects of information rigidities on propagation of nominal shocks (Lucas Jr, 1972; Woodford, 2003a; Nimark, 2008; Angeletos and La'O, 2009; Melosi, 2016; Baley and Blanco, 2019).

Among these, the closest relationship is with Maćkowiak and Wiederholt (2009) and Paciello and Wiederholt (2014). The former characterizes how firms allocate a fixed amount of attention between aggregate and idiosyncratic shocks in a monopolistic competition environment. The latter endogenizes the amount of attention as a choice variable for firms and studies optimal policy. However, the main focus of this paper is to study how *oligopolistic competition* affects both of these incentives through strategic interactions. The strategic environment considered here makes the joint distribution of idiosyncratic shocks (interpreted in the model as mistakes of firms) an endogenous object which, in turn, feeds back into the information acquisition of firms. The novel result here is that, all else equal, varying the number of competitors has quantitatively significant implications for both incentives in information acquisition and propagation of shocks.

Furthermore, the paper is also related to the literature that formalizes the incentive to learn about others' beliefs in strategic environments (Hellwig and Veldkamp, 2009). The main departure here is to relate the strength of this incentive to the number of players and understand how the two interact in a macroeconomic setting.⁵ Moreover, the fact that firms use their information about their competitors' beliefs to forecast aggregates is related to the insight in Hellwig and Venkateswaran (2009) where firms mistakenly attribute aggregate shocks to firms-specific ones.

³The message carries on to the case where firms are not myopic, but the Phillips curve does not have a closed-form expression in that case.

⁴For applications to other areas in macroeconomics, see, e.g., Luo (2008); Tutino (2013); Khaw and Zorrilla (2018) for consumption; Luo et al. (2012) for current account; Zorn (2016) for investment; Peng and Xiong (2006) for asset pricing; Mondria and Wu (2010) for home bias; and Ilut and Valchev (2017) for imperfect problem solving.

⁵In that sense, the setup of our static model here is theoretically related to the one in Denti (2018) where players can flexibly acquire correlated information.

Moreover, the dynamic model of this paper also relates to a recent literature on characterizing the solution to dynamic rational inattention models in LQG settings (Mackowiak et al., 2018; Steiner et al., 2017; Fulton, 2018; Afrouzi and Yang, 2019; Miao et al., 2020). While this literature focuses on characterizing the solution to dynamic rational inattention models for a single decision-maker, the main departure in this paper is to utilize these tools and extend them to a game-theoretic framework with strategic interactions.

Finally, the study of monetary non-neutrality in this paper is motivated by the empirical evidence on real effects of monetary shocks and aggregate price rigidities (Christiano et al., 1999; Romer and Romer, 2004).⁶ While the model is designed to create aggregate price rigidity, the assumption that firms can adjust prices every period causes the model to miss the unconditional micro price rigidity observed in the data (Bils and Klenow, 2004; Nakamura and Steinsson, 2008). This is a common shortcoming of rational inattention models in LQG settings, and can potentially be addressed by either moving beyond the LQG setup or adding nominal rigidities such as menu costs on top of oligopolistic competition and rational inattention, both of which would be natural steps forward for future research.⁷

Outline. The paper is organized as follows. Section 2 illustrates the nature of firms' information acquisition incentives in a simplified static model and derives a set of testable predictions. Section 3 relates the predictions of the model to the firm-level survey data from New Zealand. Section 4 presents the dynamic general equilibrium model and discusses the calibration strategy. Section 5 presents the results for monetary non-neutrality and propagation of shocks. Section 6 discusses robustness and alternative parameterizations. Section 7 concludes. Moreover, all the technical derivations as well as the proofs for all the propositions and corollaries are included in Appendices B and F, for the static and dynamic models respectively.

2 A Static Model

The goal of this section is to illustrate the interaction of rational inattention and oligopolistic competition within a simple static model. The model presented here is a special case of

⁶See, also, Leeper et al. (1996); Uhlig (2005); Gertler and Karadi (2015); Nakamura and Steinsson (2018).

⁷In monopolistic competition environments, both of these cases have been studied before. See, for instance, Matějka (2015) for a price-setting model without the LQG assumption that implies rigid micro prices. Also, monetary non-neutrality under menu costs has been studied extensively (Golosov and Lucas Jr, 2007; Gertler and Leahy, 2008; Nakamura and Steinsson, 2010; Midrigan, 2011; Vavra, 2014; Alvarez et al., 2016). Moreover, in a model with both oligopolistic competition and menu costs, Mongey (2018) shows that monetary non-neutrality is amplified. Finally, see also Alvarez et al. (2011) for a model with both observation and menu costs with monopolistic competition and Yang (2019) for a model of monopolistic competition with menu costs and rational inattention.

the dynamic general equilibrium model specified in Section 4. While the general dynamic model has to be solved computationally, the solution to the static case is in closed form and provides intuition for interpreting the results in later sections. In the main text, I focus on the economics of the forces at work. All informal claims in this section are formalized in Appendix B, and the proofs for propositions are included in Appendix B.8.

2.1 The Environment

There are a large number of sectors in the economy indexed by $j \in \{1, \dots, J\}$, and within every sector there are K firms. Let index j, k denote firm k in sector j . K here represents the number of firms in a specific (sub)industry that *directly* compete with one another.⁸

Firms are price setters and their profits are affected by a normally distributed fundamental shock that I denote by $q \sim \mathcal{N}(0, 1)$. For any realization of the fundamental, and a set of prices chosen by firms across the economy, $(q, p_{j,k})_{(j,k) \in J \times K}$, the losses of firm j, k in profits is given by

$$L_{j,k}((q, p_{j,k})_{j,k \in J \times K}) = (p_{j,k} - (1 - \alpha)q - \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l})^2,$$

where $\alpha \in [0, 1)$ denotes the degree of *within* industry strategic complementarity.⁹

To illustrate the importance of endogenizing information choices of firms in this environment, let us briefly consider the case of exogenous information. For an endowed information structure for the economy, aggregating the best responses of firms in pricing, we get the following expression for the aggregate price:

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

where $\overline{\mathbb{E}^{j,k}[q]}$ is the average expectation across firms of the fundamental, and $\overline{\mathbb{E}^{j,k}[p_{j,-k}]}$ is their average expectation of their *own* competitors' prices. While this equation resembles the usual result in beauty contest games, the key departure here is that the aggregate price no longer depends on the average expectation of the aggregate price across firms. Instead, it depends on the average expectation of firms about their own-industry prices.¹⁰

Therefore, in order to understand how prices are determined in the economy, one needs

⁸When asked how many direct competitors they face in their main product market, firms in New Zealand report an average of 8 (See Figure (1)).

⁹Here the fundamental q , and prices, $(p_{j,k})_{j \in J, k \in K}$, can be interpreted as log-deviations from a steady state symmetric equilibrium, which allows us to normalize their mean to zero. I micro-found this function in the dynamic model, where the quadratic loss is based on a second order approximation to the profit function of oligopolistic firms and α depends on the household's demand for their goods.

¹⁰See, for instance, Morris and Shin (2002); Angeletos and Pavan (2007) for a discussion of beauty contests with exogenous information sets, and the value of information within them.

to understand how firms form their expectations of both the fundamental as well as the prices of their competitors.

2.2 Firms' Problem

Firms make two choices. First, they choose an information structure subject to the finite amount of attention that informs them about the fundamental and the prices of their competitors. Second, they choose a pricing strategy that maps the realization of their signals to a price.

I model the information choice problem of the firms using rational inattention where arbitrarily precise information about shocks and beliefs of others are available. However, information is costly and firms have to trade off the precision of information with its cost. Subject to this cost, firms choose their information set to maximize their ex-ante payoffs.

Formally, I assume that there is a set of available signals in the economy, denoted by \mathcal{S} , that is *rich* insofar that it allows firms to freely choose the joint distribution between their own signals, their competitors' signals and the fundamental. In particular, players are allowed to choose distributions that imply correlated beliefs, even conditional on the exogenous shocks. This requires a careful definition of the strategies along with a *rich* set of available information that allows for such strategies. A formal definition and characterization of a such an information set is provided in Appendix B.2.¹¹

Moreover, a byproduct of assuming a rich set of available information is that firms always choose to observe only one signal. Lemma B.4 in Appendix B.4 provides a formal proof of this claim.¹² Intuitively, the result hinges on the fact that information is only valued by firms insofar it allows them to choose a better price. Therefore, under the optimal information choice, a firm's price should be a sufficient statistic for its signal(s). Richness is a necessary condition for the existence of such signals.

Therefore, a pure strategy for firm j, k is to choose a signal $S_{j,k} \in \mathcal{S}$, and a pricing strategy $p_{j,k} : S_{j,k} \rightarrow \mathbb{R}$. Given a strategy profile for others, $(S_{l,m} \in \mathcal{S})_{(l,m) \neq (j,k)}$, firm j, k 's problem is

$$\min_{S_{j,k} \in \mathcal{S}} \mathbb{E} \left[\min_{p_{j,k} : S_{j,k} \rightarrow \mathbb{R}} \mathbb{E} \left[\left(p_{j,k}(S_{j,k}) - (1 - \alpha)q - \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l}(S_{j,l}) \right)^2 \mid S_{j,k} \right] \right] \quad (2)$$

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

¹¹My definition of a rich information set corresponds to the concept of flexibility in information acquisition in Denti (2018).

¹²See also Steiner et al. (2017); Mackowiak et al. (2018); Afrouzi and Yang (2019) for proofs of similar lemmas in single decision making problems.

where $\mathcal{I}(S_{j,k}; (q, p_{l,m}(S_{l,m}))_{(l,m) \neq (j,k)})$ measures the amount of information that the firm's signal reveals about the fundamental shock and the prices of other firms in the economy.¹³ This constraint simply requires that a firm cannot know more than κ bits about the fundamental q and the signals that others have chosen in \mathcal{S} . In the dynamic model, we also endogenize κ as a choice variable for firms.

Definition 1. A pure strategy Gaussian equilibrium for this economy is a strategy profile $(S_{j,k} \in \mathcal{S}, p_{j,k} : S_{j,k} \rightarrow \mathbb{R})_{(j,k) \in J \times K}$ such that $\forall (j,k) \in J \times K$, $(S_{l,m}, p_{l,m})_{(l,m) \neq (j,k)}$ solves j, k 's problem as stated in Equation (2).

It is shown in Appendix B that the equilibrium is unique in the joint distribution of prices for firms. This result allows us to abstract away from characterizing the underlying signals and directly focus on how firms' prices are related to one another. Let $p_{j,k}$ be the price that firm j, k charges in the equilibrium. Then the unique joint distribution of prices and the fundamental shock are characterized by

$$\begin{aligned}
 p_{j,k} &= \lambda \left((1 - \alpha)q + \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l} \right) + z_{j,k} \\
 z_{j,k} &\perp (q, S_{m,l})_{(m,l) \neq (j,k)} \\
 \mathbb{E}[z_{j,k}] &= 0, \quad \text{Var}(z_{j,k}) = \lambda(1 - \lambda) \text{Var} \left((1 - \alpha)q + \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l} \right)
 \end{aligned} \tag{3}$$

Where $\lambda \equiv 1 - 2^{-2\kappa}$ and $z_{j,k}$ is noise in prices introduced by rational inattention. Appendix B.6 shows that this system of equations uniquely pins down the equilibrium distribution of prices and the fundamental shock. Two observations immediately follow: larger capacity, κ , increases the covariance of prices with the fundamental and decreases the variance of the rational inattention noise. In particular, when $\kappa \rightarrow \infty$, λ approaches 1, the noise disappears and $p_{j,k} = q, \forall j, k$. The rest of this section unpacks the properties of this solution and studies its economic implications for given levels of κ , α and K . Afterwards, in the dynamic model, I allow firms to endogenously choose κ , micro-found α through a demand system and introduce heterogeneity in K that I then calibrate to the survey data.

¹³ $\mathcal{I}(\cdot; \cdot)$ is Shannon's mutual information function. In this paper, I focus on Gaussian random variables, in which case $\mathcal{I}(X; Y) = \frac{1}{2} \log_2(\det(\text{var}(X))) - \frac{1}{2} \log_2(\det(\text{var}(X|Y)))$. The Gaussian nature of the information structure is self-consistent in the equilibrium. When a firms' opponents choose Gaussian signals, under the quadratic loss it is also optimal for the firm to choose a Gaussian signal. See Cover and Thomas (2012) for optimality of Gaussian signals under quadratic objectives with Gaussian fundamentals.

2.3 Attention Allocation

The finite attention of a firm caused by a finite κ implies that this price cannot be fully revealing of the fundamental and contains the firm's perception noise caused by inattention. This noise acts as a mistake in observing the fundamental and affects the price of the firm and, accordingly, the profits of its competitors.

Definition 2. A *mistake* is a part of a firm's price that is unpredictable by the fundamentals of the economy.

Thus, any firm's price can be decomposed into the part that is correlated with the fundamental and the part that is orthogonal to it:

$$p_{j,k} = \delta q + v_{j,k}, \quad v_{j,k} \perp q, \quad \delta \in \mathbb{R}.$$

The vector $(v_{j,k})_{j,k \in J \times K}$, therefore, contains the *mistakes* of all firms in pricing, with their joint distribution endogenously determined in the equilibrium.

It is important to mention that these mistakes need not be independent across firms. In fact, by endogenizing the information choices of firms, one of the objectives here is to understand how the mistakes of different firms relate to one another in the equilibrium, or intuitively how much managers of competing firms *attend* to the mistakes of their rivals and incorporate them in their own prices. Moreover, the coefficient δ , which determines the degree to which prices covary with the fundamental of the economy, is also an equilibrium object. Our goal is to understand how δ and the joint distribution of mistakes rely on the underlying parameters of the model: α , K and κ .

Definition 3. The *amount of attention* that a firm pays to a random variable is the mutual information between their set of signals and that random variable. Moreover, we say a firm *knows more about X than Y* if it pays more attention to X than Y.

In the static model, the amount of attention is directly linked to the absolute value of the correlation between a firm's signal and the random variable to which the firm is paying attention.¹⁴ Appendix B.7 shows that when others play a strategy in which $\frac{1}{K-1} \sum_{l \neq k} p_{j,l} = \delta q + v_{j,-k}$, the attention problem of firm j, k reduces to choosing the correlation of their signal with the fundamental and the mistakes of others:

$$\begin{aligned} \max_{\rho_q \geq 0, \rho_v \geq 0} \quad & \rho_q + \frac{\alpha \sigma_v}{1 - \alpha(1 - \delta)} \rho_v \quad , \\ \text{s.t.} \quad & \rho_q^2 + \rho_v^2 \leq \lambda \equiv 1 - 2^{-2\kappa} \quad . \end{aligned}$$

¹⁴For two normal random variables X and Y , let $\mathcal{I}(X, Y)$ denote Shannon's mutual information between the two. Then $\mathcal{I}(X, Y) = -\frac{1}{2} \log_2(1 - \rho_{X,Y}^2)$ where $\rho_{X,Y}$ is the correlation between X and Y . Notice that $\mathcal{I}(X, Y)$ is increasing in $\rho_{X,Y}^2$.

Here $\sigma_v \equiv \text{var}(v_{j,-k})^{\frac{1}{2}}$ is the standard deviation of the average mistakes of j, k 's competitors, ρ_q is the correlation of the firm's signal with the fundamental, and ρ_v is its correlation with the average mistake of its competitors.

The following proposition states the properties of the equilibrium. The closed form solutions and derivations are included in the proof in Appendix B.

Proposition 1. *In equilibrium,*

1. *Firms pay strictly positive attention to the mistakes of their competitors ($\rho_v^* > 0$) if $\alpha > 0$ and K is finite.*
2. *Firms' knowledge of the fundamental increases in the number of their competitors and decreases in the degree of strategic complementarity:*

$$\frac{\partial}{\partial K} \rho_q^* > 0, \quad \frac{\partial}{\partial \alpha} \rho_q^* < 0.$$

3. *Firms do not pay attention to mistakes of those in other industries: $\forall (j, k), (l, m)$, if $j \neq l$, $p_{j,k} \perp p_{l,m} | q$.*

The first part of Proposition 1 follows from the fact that firms are affected by the mistakes of their competitors and find it optimal to pay strictly positive attention to them. However, since mistakes are orthogonal to the fundamental, any attention to others' mistakes has to be traded off with attention to the fundamental.

The second part of Proposition 1 shows how the incentive to track others' mistakes is affected by the degree of strategic complementarity and the number of a firm's competitors. α is the underlying parameter that relates the payoff of a firm to the mistakes of its competitors. The larger is α , the more the payoffs of a firm depends on the mistakes of its competitors. Accordingly, the firm finds it more in their interest to track those mistakes. This illustrates the importance of micro-founding these strategic complementarities, which is one of the main objectives of the model in Section 4.

Moreover, the direct effect of K is captured by the presence of σ_v in the objective of the firms and the fact that firms are only affected by the average of their rivals' mistakes. This variance gets smaller as the number of a firm's competitors increases and goes to zero as $K \rightarrow \infty$.¹⁵ Therefore, the larger the number of a firm's competitors, the more their mistakes "wash out". This allows a more competitive firm to substitute away from paying attention to others' mistakes and pay more attention to the fundamental shock.

¹⁵This is an equilibrium outcome as σ_v is determined by the endogenous choices of firms. To see how this emerges in the equilibrium, notice that if $K \rightarrow \infty$, a firm has no incentive to pay attention to others' mistakes if they are independent as the law of large numbers would imply $\sigma_v = 0$. Since incentives are symmetric, in the equilibrium all firms prefer to have independent mistakes, implying that $\sigma_v = 0$.

Finally, Appendix B shows that, in equilibrium, the covariance of aggregate price and the fundamental shock is given by

$$\delta = \frac{\lambda - \alpha\lambda}{1 - \alpha\lambda}.$$

This implies that the degree to which prices covary with the fundamental in an industry increases with the capacity of processing information and decreases with the degree of strategic complementarity. While it is the case that this covariance is independent of the number of firms, this is not a robust feature and goes away in dynamics where strategic complementarities are micro-founded and beliefs are dynamic.

2.4 Equilibrium Prices and Expectations

In conventional models, price setters' expectations of the aggregate price is a crucial element of their pricing decisions. However, the empirical evidence on firms' expectations about aggregate inflation suggests that this link is not present in the data, and there is a disconnect between firms' prices and their expectations of aggregate inflation. This simple model, however, predicts a different relationship between prices and expectations. Here, firms' prices are related to their expectations of their competitors' prices, and the aggregate price is related to an average of those expectations:

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

The following proposition shows that this model creates a wedge between prices and aggregate expectations of firms about the aggregate price.

Proposition 2. *In equilibrium, the aggregate price co-moves more with the average expectations from own-industry prices than average expectations of the aggregate price itself, meaning that*

$$\text{cov}(p, \overline{\mathbb{E}^{j,k}[p_{j,-k}]}) > \text{cov}(p, \overline{\mathbb{E}^{j,k}[p]}).$$

Moreover, the two converge to each other as $K \rightarrow \infty$.

Therefore, what firms know about the prices of their competitors matters more for the determination of the aggregate price than what they know about the aggregate price itself. This result also holds in the dynamic model in the sense that inflation is driven more by the expectations of industry price changes than by the expectations over inflation itself. The following corollary shows that the realized price is also closer to the average own-industry price expectations than the average expectation of the aggregate price.

Corollary 1. *In equilibrium, the realized price is closer in absolute value to the average expectations from own-industry prices than the average expectation of the aggregate price itself.*

$$|p - \overline{\mathbb{E}^{j,k}[p_{j,-k}]}| < |p - \overline{\mathbb{E}^{j,k}[p]}|$$

The intuition behind these results relies solely on the incentives of firms in paying attention to the mistakes of their competitors. In equilibrium, the signals that firms observe are more informative of their own industry prices than the aggregate economy:

$$S_{j,k} = \underbrace{\overbrace{p}^{\text{covaries with aggregate price}}}_{\text{covaries with industry prices}} + u_j + e_{j,k}$$

where we have decomposed the mistake of firm j, k as $v_{j,k} = u_j + e_{j,k}$, where $u_j \perp p$ is the common mistake in industry j and $e_{j,k}$ is the independent part of firm j, k 's mistake. The fact that $\text{var}(u_j) \neq 0$ by Proposition 1 implies that the firm is more informed in predicting its own industry price changes than the aggregate price, and the two would become the same only if there was no coordination within industries in information acquisition, which happens when $K \rightarrow \infty$.

This result, along with its counterpart in the dynamic model, shows how stable inflation can be an equilibrium outcome even when firms' expectations of that inflation are ill-informed. What firms need to know in terms of figuring out their optimal price is a combination of the fundamental q and their own industry price changes. While the aggregate price will be correlated with both of these objects, it does not by itself play an important role in firms' profits and firms do not need to directly learn about it.

3 Model Predictions and the Survey Data

The goal of this section is twofold: first, to provide evidence for the main assumptions of the model and, second, to test the main predictions of the simple model against data. To do so, I use a unique quantitative survey of firms' expectations from New Zealand, which is comprehensively discussed in [Kumar, Afrouzi, Coibion and Gorodnichenko \(2015\)](#) and [Coibion, Gorodnichenko and Kumar \(2018\)](#), to assess the predictions of the model in the previous section. The survey was conducted in multiple waves among a random sample of firms in New Zealand with broad sectoral coverage.

The new empirical contribution in this paper relative to the previous papers that have used this data is that I (1) implement and utilize a new question in the survey to back-out the degree of strategic complementarity for firms, and (2) document that firms with more

competitors have more certain posteriors about the aggregate inflation.

3.1 Number of Competitors and Strategic Complementarity

Two assumptions are crucial for the results of the model in the previous section: the finiteness of a firm’s competitors and the existence of micro-level strategic complementarities. Two questions in the survey directly measure these for every firm within the sample and quantify these assumptions. The first question asks firms

“How many direct competitors does this firm face in its main product line?”

Figure (1) shows the distribution of firms’ responses to this question. Columns (1) and (2) in Table (1) show that the average firm in the sample faces only eight competitors with 45% of firms reporting that they face six or fewer competitors. A breakdown of firms’ answers from different industries shows that this average is fairly uniform across them.

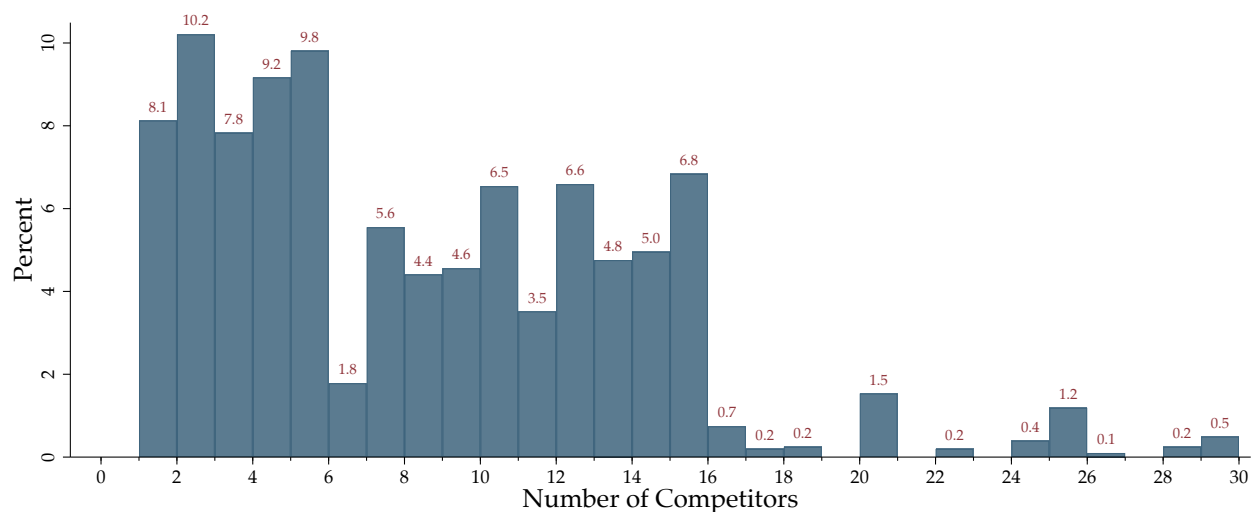


Figure 1: Distribution of the Number of Competitors

Notes: the figure presents the the distribution of the number of competitors that firms report they face in their direct product market in the survey data from New Zealand. The numbers over bars denote the percentage of firms within the corresponding bin. Firms with more than 30 competitors are dropped (only less than 1 percent of firms report they have more than 30 competitors, with a max of 42).

We also implemented a question in the survey to measure the degree of micro-level strategic complementarity. This has been a challenging parameter to estimate in the literature due to major endogeneity concerns: it is rarely possible to find exogenous variations in

the prices of a firms’ competitors that are not correlated with aggregates or the firm’s own costs. To bypass this issue, I rely on the following hypothetical question to measure the degree of strategic complementarity:

- “Suppose that you get news that the general level of prices went up by 10% in the economy:
- By what percentage do you think your competitors would raise their prices on average?
 - By what percentage would your firm raise its price on average?
 - By what percentage would your firm raise its price if your competitors did not change their price at all in response to this news?”

The question proposes a change in the firms’ environment that is coming through aggregate variables, which affects both their costs and those of their competitors. The question then measures three different quantities that allow me to disentangle the degree of strategic complementarity for the firm:

$$p_{j,k} = \underbrace{(1 - \alpha)\mathbb{E}^{j,k}[q]}_{\text{Answer to c.}} + \underbrace{\alpha\mathbb{E}^{j,k}[p_{j,-k}]}_{\text{Answer to a.}}. \quad (4)$$

The average α implied by the responses of firms to this question is 0.82 and uniform across different industries, as reported in columns (3) and (4) of Table (1). The usual calibration for the strategic complementarity in the U.S. in monopolistic competition models is 0.9 which is slightly larger than what I estimate here.¹⁶

Table 1: Number of Competitors and Degree of Strategic Complementarity

	Number of Competitors		Strategic Complementarity	
	(1)	(2)	(3)	(4)
Constant	8.449 (0.113)	8.367 (0.198)	0.817 (0.008)	0.795 (0.015)
Manufacturing		-		-
Construction		-1.285 (0.425)		0.063 (0.032)
Trade		0.183 (0.334)		0.010 (0.026)
Services		0.319 (0.287)		0.031 (0.021)
Observations	3072	2667	2824	2445

Standard errors in parentheses

Notes: the table presents statistics for the number of competitors and the degree of strategic complementarity in the survey data from New Zealand. Columns (1) and (2) report the average number of competitors that firms report they face in their main product market. Columns (3) and (4) show the coefficient for the degree of strategic complementarity from Equation (4).

¹⁶See, for instance, [Mankiw and Reis \(2002\)](#); [Woodford \(2003b\)](#).

3.2 Knowledge about Industry versus Aggregate Inflation

One of the main predictions of the model is that firms are more aware of their competitors' price changes than the aggregate price.

In the fourth wave of the survey, conducted in the last quarter of 2014, firms were asked to provide their nowcasts of both industry and aggregate yearly inflation.

Table 2: Size of Firms' Nowcast Errors

	<i>Observations</i>	<i>Industry inflation</i>		<i>Aggregate inflation</i>	
		<i>mean</i>	<i>std</i>	<i>mean</i>	<i>std</i>
<i>Industry</i>	(1)	(2)	(3)	(4)	(5)
Construction	52	0.75	0.54	3.95	1.95
Manufacturing	363	1.43	1.72	2.55	2.04
Financial Services	352	1.51	1.51	4.23	1.73
Trade	302	0.63	0.90	2.31	1.93
Total	1,069	1.20	1.49	3.11	2.09

Notes: the table reports the size of firms' nowcast errors in perceiving aggregate inflation versus industry inflation for the 12 months ending in December 2014.

Table (2) reports the size of firms' nowcast errors in perceiving these two inflation rates.¹⁷ The average absolute nowcast error across firms about their own industry inflation is 1.2 percentage points, a magnitude that is considerably lower than the average absolute nowcast error about aggregate inflation, 3.1 percentage points.

Furthermore, Figure (2) shows that in addition to this striking difference in the averages, the distributions of these nowcast errors are skewed in opposite directions: for nearly two-thirds of firms, their nowcast error of the aggregate inflation is larger than the mean error, while the reverse is true in the case of industry inflation.

3.3 Uncertainty about Inflation versus Number of Competitors

In the sixth wave of the survey, conducted in 2016, firms were asked to report the distribution of their beliefs about both aggregate and their industry inflation through the following two similarly worded questions:

"Please assign probabilities (from 0-100) to the following ranges of overall price changes in [the economy] [your industry] over the next 12 months for New Zealand."

¹⁷Nowcast errors for industry inflation are measured as the distance between firms' nowcast and the realized inflation in their industry.

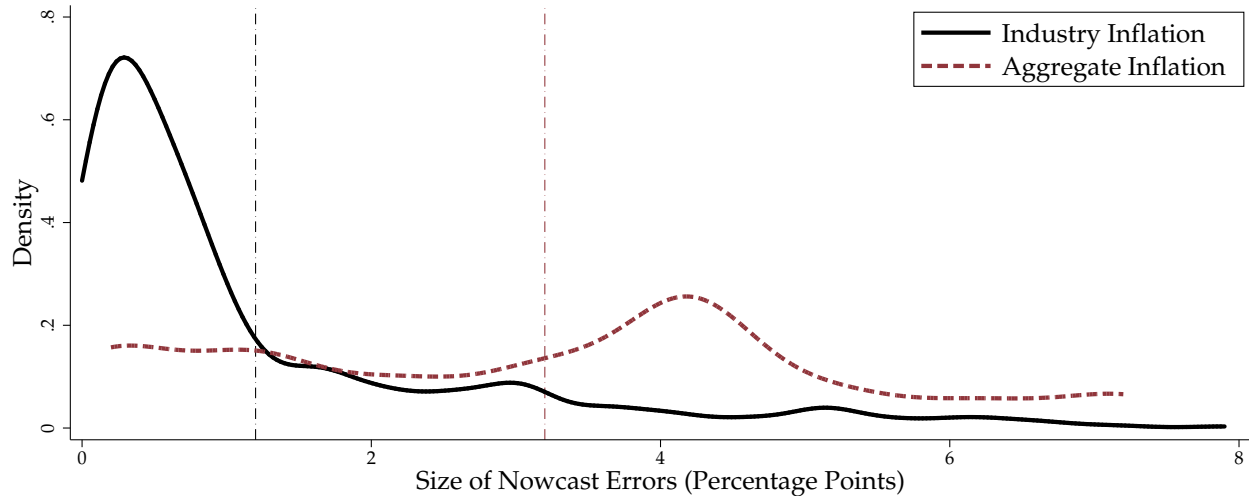


Figure 2: Distributions of the Size of Firms' Nowcast Errors

Notes: the figure presents the distribution of the size of firms' errors in perceiving the aggregate and their industry inflation in New Zealand. The dashed vertical lines denote the means of these distributions.

For both questions, firms were provided with an identical set of bins to which they assigned their subjective probabilities.¹⁸

Proposition 1 predicts that knowledge about the aggregate price should be increasing in the number of a firm's competitors. This is a unique feature of the oligopolistic rational inattention model and is a testable prediction. To test this prediction, I run the following regression:

$$\log(\sigma_i^\pi) = \beta_0 + \beta_1 \log(K_i) + \epsilon_i, \quad (5)$$

where σ_i^π is firm i 's subjective uncertainty about the aggregate inflation, and K_i is the number of competitors that they report in their main product market. The model's prediction translates to the null hypothesis that $\beta_1 < 0$. Table (3) reports the result of this regression, and shows that this is indeed the case. This result is also robust to including firm controls such as firms' age and employment as well as industry fixed effects.

¹⁸These were two separate questions in the survey that I have combined here in order to avoid repetition. The assigned bins varied from -25 percent to 25 percent with 5 percent increments. The wide range is provided because firms are highly uncertain about inflation and assign positive probabilities to high inflation rates. The large negative magnitudes were also provided to avoid priming concerns.

Table 3: Subjective Uncertainty of Firms and the Number of Competitors.

	$\log(\sigma_i^\pi)$	
	(1)	(2)
log(#competitors)	-0.116 (0.012)	-0.113 (0.013)
Firm controls and FEs	No	Yes
Observations	1,662	1,552

Robust standard errors in parentheses

Notes: the table reports the result of regressing the log standard deviation of firms' reported distribution for their forecast of aggregate inflation on the log of their number of competitors as well as a set of firm controls (age, size measured by employment, and fixed effects for construction, manufacturing, financial services and trade industries).

The significance of this coefficient in explaining firms' uncertainty about aggregates is an observation that is not reconcilable with full information rational expectation models or, to the best of my knowledge, any other macroeconomic model of information rigidity prior to this paper, and indicates the importance of strategic incentives in how much firms pay attention to aggregate variables in the economy.

4 A Micro-founded Dynamic Model

The goal of this section is to extend the simple static model of Section 2 to a dynamic general equilibrium model to quantitatively analyze the effects of firms' strategic incentives in information acquisition for the propagation of monetary policy shocks to aggregate output and inflation.

In particular, the model in this section improves on the static model by (1) micro-founding the loss function and micro-level strategic complementarities as a function of a representative household's demand for different varieties of goods (2) endogenizing the choice of information processing capacity on the part of firms, and (3) considering the dynamic incentives of firms in information acquisition in addition to their strategic incentives (dynamically inattentive firms realize that information has a continuation value and choose their information accordingly).

All the derivations as well as the proofs for the propositions regarding the dynamic model are included in Appendix F.

4.1 Environment

4.1.1 Households

There is a large variety of goods produced in the economy. In particular, the economy consists of a large number of sectors, $j \in J \equiv \{1, \dots, J\}$; and each sector j consists of $K_j \geq 2$ firms that produce weakly substitutable goods. Here, K_j is drawn from an exogenous distribution \mathcal{K} (that I will, later on, calibrate to the distribution of the number of competitors in the data). The representative household takes the nominal prices of these goods as given and forms a demand over the product of each firm in the economy. In particular, the aggregate time t consumption of the household is

$$C_t \equiv \prod_{j \in J} C_{j,t}^{J^{-1}}, \quad (6)$$

where $C_{j,t}$ is the composite demand of the household for the goods produced in sector j and is determined by a CES aggregation of within sector goods with an elasticity of substitution $\eta > 1$.¹⁹ Equation (6) denotes that the aggregate consumption of the household is Cobb-Douglas in the composite goods of sectors. Finally, I assume that the representative household has full information rational expectations.²⁰

Therefore, the representative household's problem is

$$\begin{aligned} & \max_{((C_{j,k,t})_{(j,k) \in J \times K}, C_t, L_t, B_t)_{t=0}^{\infty}} \mathbb{E}_0^f \sum_{t=0}^{\infty} \beta^t [\log(C_t) - L_t] \\ \text{s.t.} \quad & \sum_{j,k} P_{j,k,t} C_{j,k,t} + B_t \leq W_t L_t + (1 + i_{t-1}) B_{t-1} + \sum_{j,k} \Pi_{j,k,t} - T \\ & C_t = \prod_{j \in J} \left[\left(K_j^{-1} \sum_{k \in K_j} C_{j,k,t}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \right]^{J^{-1}} \end{aligned} \quad (7)$$

where $\mathbb{E}_t^f[\cdot]$ is the full information rational expectations operator at time t , L_t is the labor supply of the household, B_t is their demand for nominal bonds, W_t is the nominal wage, i_t is the net nominal interest rate, $\Pi_{j,k,t}$ denotes the profit of firm j, k at time t , and T is a constant lump sum tax that is used by the government to finance a hiring subsidy for firms

¹⁹A more general aggregator can be considered here – see, e.g., [Rotemberg and Woodford \(1992\)](#). I derive the implied demand under a general form for this aggregator function in Appendix D. Another specific case is the Kimball aggregator which I discuss in Appendix E.

²⁰Since the main purpose of this paper is to study the effects of rational inattention under imperfect competition among firms, I assume that households are fully informed about prices and wages. This is a common assumption in the literature – see e.g. [Melosi \(2016\)](#) – and simplifies the household side of the economy as a natural first step in separating the implications of rational inattention for households versus firms.

in order to eliminate any long-run inefficiencies of imperfect competition.

The CES aggregator within sector goods leads to the following demand function for the product of firm j, k .

$$C_{j,k,t} = Q_t \mathcal{D}(P_{j,k,t}; P_{j,-k,t}) \quad (8)$$

$$\mathcal{D}(P_{j,k,t}; P_{j,-k,t}) \equiv \frac{P_{j,k,t}^{-\eta}}{\sum_{l \in K_j} P_{j,l,t}^{1-\eta}}$$

where $Q_t \equiv P_t C_t$ is the nominal aggregate demand (P_t is the price of the aggregate consumption bundle C_t), $P_{j,k,t}$ is firm j, k 's price at t , and $P_{j,-k,t}$ is the vector of other firms' prices in sector j . Furthermore, the household's intertemporal Euler and labor supply equations are given by:

$$W_t = Q_t, \quad 1 = \beta(1 + i_t) \mathbb{E}_t^f \left[\frac{Q_t}{Q_{t+1}} \right].$$

The log-utility implies that the intertemporal Euler equation simply relates the level of nominal interest rate to the expected inverse growth of the aggregate demand, and the linear disutility of labor implies that nominal wage is proportional to the nominal demand.²¹

4.1.2 Firms

Firms are *rationally inattentive*. At the beginning of each period t , they take their initial information set as given and choose an arbitrary number of signals from a rich set of *available signals*, \mathcal{S}^t , subject to an information processing constraint.²² In contrast to the static model where I assumed this capacity was exogenous, here I assume firms can produce this capacity by hiring labor from the competitive labor market. The production function for this capacity is linear in labor and is given by $\kappa_{j,k,t} = \omega^{-1} L_{j,k,t}^i$ where $L_{j,k,t}^i$ is the labor demand of the firm for producing information processing capacity and ω^{-1} captures the productivity of labor in producing it. After firms choose the joint distribution of their new signals with the fundamental shocks, all new shocks and new signals are realized. Firms then form their new information set by adding their new observed signals to their last period information set and choose their prices based on that.

Since my main objective is to examine the real effects of monetary policy through endogenous information acquisition of these firms, I abstract away from other sources of mon-

²¹The linear disutility in labor is a common assumption in the models of monetary non-neutrality (for instance, see Golosov and Lucas Jr (2007)) which eliminates the source of *across* sector strategic complementarity from the household side. I use this assumption to the same end in order to mainly focus on micro-founding *within* sector strategic complementarities.

²²See Appendix C for the formal specification of \mathcal{S}^t .

etary non-neutrality, and in particular, assume that prices are perfectly flexible.²³

After setting their prices every period, firms' demands are realized and they hire labor from the competitive labor market to produce with a production function that has decreasing returns in labor; $Y_{j,k,t} = (L_{j,k,t}^p)^{\frac{1}{1+\gamma}}$. Here, $\gamma = 0$ corresponds to constant returns to scale and positive γ captures the degree of decreasing returns to scale in labor.

Formally, a strategy for firm j, k is to choose a capacity for processing information conditional on their initial information set at any time, $\kappa_{j,k,t} = \omega^{-1} L_{j,k,t}^i : S_{j,k}^{t-1} \rightarrow \mathbb{R}_+$, a set of signals to observe, $S_{j,k,t} \subset \mathcal{S}^t$, and a pricing strategy that maps its information set to their optimal actions, $P_{j,k,t} : S_{j,k}^t \rightarrow \mathbb{R}$, where $S_{j,k}^t = \{S_{j,k,\tau}\}_{\tau=0}^t$ is the firm's information set at time t . Firms then hire enough labor for their good production to satisfy demand.²⁴ Given a strategy for all the other firms in the economy, firm j, k 's problem is to maximize the net present value of their lifetime profits given an initial information set that they inherit from the previous period:

$$\begin{aligned}
 & \max_{\{S_{j,k,t} \subset \mathcal{S}^t, P_{j,k,t}(S_{j,k}^t), L_{j,k,t}^i(S_{j,k}^{t-1})\}_{t \geq 0}} & (9) \\
 & \mathbb{E} \left[\sum_{t=0}^{\infty} \underbrace{\beta^t Q_t^{-1}}_{\text{discount factor}} \left(\underbrace{P_{j,k,t} Y_{j,k,t}^d}_{\text{revenue}} - \underbrace{(1 - \bar{s}_j) W_t ((Y_{j,k,t}^d)^{1+\gamma} + L_{j,k,t}^i)}_{\text{costs of goods and information production}} \right) \middle| S_{j,k}^{-1} \right] \\
 \text{s.t. } & Y_{j,k,t}^d = Q_t \mathcal{D}(P_{j,k,t}; P_{j,-k,t}) & (\text{demand}) \\
 & \kappa_{j,k,t} = \omega^{-1} L_{j,k,t}^i & (\text{information capacity production technology}) \\
 & \mathcal{I}(S_{j,k,t}, (Q_\tau, P_{l,m,\tau}(S_{l,m}^\tau))_{0 \leq \tau \leq t}^{(l,m) \neq (j,k)} | S_{j,k}^{t-1}) \leq \kappa_{j,k,t} & (\text{information processing constraint}) \\
 & S_{j,k}^t = S_{j,k}^{t-1} \cup S_{j,k,t}, \quad S_{j,k}^{-1} \text{ given.} & (\text{evolution of the information set})
 \end{aligned}$$

where the information processing constraint requires that the amount of information that a firm can add to its information set about the state of the economy at a given time is bounded above by its produced capacity $\kappa_{j,k,t}$. Here, the function $\mathcal{I}(\cdot, \cdot)$ is Shannon's mutual information function, which measures the reduction in conditional entropy experienced by the firm across two consecutive periods.²⁵ Moreover, \bar{s}_j denotes a constant hiring subsidy to firms in sector j that eliminates the steady state inefficiencies from imperfect competition

²³There is also a new growing literature that argues information rigidities are more consistent with certain aspects of the pricing behavior of firms rather than Calvo pricing or menu cost models. For instance, see, [Stevens \(2020\)](#); [Khaw, Stevens and Woodford \(2017\)](#).

²⁴This is a ubiquitous assumption in the literature with sticky prices, which rules out temporary shutdowns of production by firms due to negative profits induced by suboptimal prices. See e.g. [Woodford \(2003b\)](#); [Goloso and Lucas Jr \(2007\)](#). Formally, this assumption requires that supply *has* to be equal to demand.

²⁵See Appendix B.1 for the formal specification of Shannon's mutual information function.

and implements the optimal level of output in that steady state.

4.1.3 Monetary Policy and General Equilibrium

I assume that the monetary policy is set in terms of the growth of nominal aggregate demand. This is justified by the household's intertemporal Euler equation as it establishes a direct relationship between nominal rates and the expected growth in nominal demand and is a standard approach for modeling monetary policy in models of monetary non-neutrality.²⁶ Following the literature, I particularly assume that this growth rate is an AR(1) process with a persistence of ρ :

$$\log\left(\frac{Q_t}{Q_{t-1}}\right) = \rho \log\left(\frac{Q_{t-1}}{Q_{t-2}}\right) + u_t. \quad (10)$$

Definition 4. A general equilibrium for the economy is an allocation for the household,

$$\Omega^H \equiv \{(C_{j,k,t})_{j \in J, k \in K_j}, L_t^s, B_t\}_{t=0}^\infty,$$

a strategy profile for firms given an initial set of signals

$$\Omega^F \equiv \{(S_{j,k,t} \subset \mathcal{S}^t, P_{j,k,t}, L_{j,k,t}^i, L_{j,k,t}^p, Y_{j,k,t}^d)_{t=0}^\infty\}_{j \in J, k \in K_j} \cup \{S_{j,k}^{-1}\}_{j \in J, k \in K_j},$$

and a set of prices $\{i_t, P_t, W_t\}_{t=0}^\infty$ such that

1. Households: given prices and Ω^F , the household's allocation solves their problem as specified in Equation (7).
2. Firms: given prices and Ω^H , and the implied labor supply and output demand curves, no firm has an incentive to deviate from Ω^F .
3. Monetary Policy: given prices, Ω^F and Ω^H , $\{Q_t \equiv P_t C_t\}_{t=0}^\infty$ satisfies the monetary policy rule specified in Equation (10).
4. Markets clear:

$$\text{Goods Markets:} \quad C_{j,k,t} = Y_{j,k,t}^d, \quad \forall j \in J, k \in K_j,$$

$$\text{Labor Markets:} \quad \sum_{j \in J, k \in K_j} (L_{j,k,t}^p + L_{j,k,t}^i) = L_t^s \quad .$$

²⁶See, for instance, [Mankiw and Reis \(2002\)](#); [Woodford \(2003a\)](#); [Golosov and Lucas Jr \(2007\)](#); [Nakamura and Steinsson \(2010\)](#).

4.2 Sources of Strategic Complementarity

Strategic complementarities in pricing are at the core of this paper's focus on understanding how firms allocate their attention across aggregate variables and the prices of their competitors. Therefore, a brief discussion of the sources of strategic complementarities in the model is essential.

There are two sources of strategic complementarity in the model: (1) decreasing returns to scale in labor ($\gamma > 0$) and (2) sensitivity of optimal markups to relative prices. Complementarities due to decreasing returns to scale are not specific to oligopolistic environments and are commonly used in monetary models. They exist because firms' marginal costs are sensitive to their levels of production – which in turn depends on their relative prices. I assume decreasing returns to scale mainly for calibration purposes.²⁷

However, the sensitivity of optimal markups to relative prices under CES demand is only a source for strategic complementarity when firms are oligopolistic. Contrary to models of monopolistic competition where a constant elasticity of substitution across varieties implies a constant markup for firms over their marginal costs, an oligopolistic environment relates these markups to firms' relative prices. This is because granularity of firms in an oligopoly implies that any change in a firm's price influences the distribution of demand across their competitors. Accordingly, demand elasticities for firms within an oligopoly depends on the relative prices of all those firms and is no longer a constant. A look at the best response of a firm to a particular realization of $P_{j,-k,t}$ and Q_t manifests this relationship:

$$P_{j,k,t}^* = \underbrace{\mu(P_{j,k,t}^*, P_{j,-k,t})}_{\text{optimal markup}} \times \underbrace{(1 - \bar{s}_j)(1 + \gamma)Q_t^{1+\gamma} \mathcal{D}(P_{j,k,t}^*, P_{j,-k,t})^\gamma}_{\text{marginal cost}} \quad (11)$$

where $P_{j,k,t}^*$ is the implied optimal price given Q_t and the vector of others prices $P_{j,-k,t}$ and the optimal markup has the familiar expression in terms of the elasticity of a firm's demand, $\mu(P_{j,k,t}^*, P_{j,-k,t}) \equiv \frac{\varepsilon_D(P_{j,k,t}^*, P_{j,-k,t})}{\varepsilon_D(P_{j,k,t}^*, P_{j,-k,t}) - 1}$. Here, $\varepsilon_D(P_{j,k,t}, P_{j,-k,t}) \equiv -\frac{\partial Y_{j,k,t}}{\partial P_{j,k,t}} \frac{P_{j,k,t}}{Y_{j,k,t}}$ is firm j, k 's elasticity of demand with respect to its own price. Following [Atkeson and Burstein \(2008\)](#), it is informative to write these elasticities in terms of a firm's market share within its own sector:

$$\varepsilon_D(P_{j,k,t}, P_{j,-k,t}) = \eta - (\eta - 1)m_{j,k,t} \quad (12)$$

$$m_{j,k,t} \equiv \frac{P_{j,k,t} Y_{j,k,t}^d}{\sum_{l \in K_j} P_{j,l,t} Y_{j,l,t}^d} \quad (13)$$

An immediate observation is that the *level* of optimal markups increase in a firm's market

²⁷See [Woodford \(2003b\)](#) or [Galí \(2015\)](#) for discussions and applications of this channel in generating strategic complementarities.

share and converges to the monopolistic competition markup when this market share goes to zero:

$$\mu(P_{j,k,t}^*, P_{j,-k,t}) = \frac{\eta}{\eta - 1} + \frac{1}{\eta - 1} \frac{m_{j,k,t}}{1 - m_{j,k,t}} \quad (14)$$

Moreover, given the definition of market shares, one can derive the degree of strategic complementarity for a given set of prices by differentiating the best response of the firm. For the special case when there are constant returns to scale ($\gamma = 0$) so that sensitivity of markups is the only source of complementarity, this adopts a clear representation in terms of the market shares:

$$\frac{dP_{j,k,t}^*}{P_{j,k,t}^*} \Big|_{\gamma=0} = \frac{dQ_t}{Q_t} + \underbrace{(1 - \eta^{-1})m_{j,k,t}}_{\text{strategic complementarity}} \left(\underbrace{\frac{\sum_{l \neq k} m_{j,l,t} dP_{j,l,t} / P_{j,l,t}}{\sum_{l \neq k} m_{j,l,t}}}_{\text{average price-change of others}} - \underbrace{\frac{dQ_t}{Q_t}}_{\text{change in wage}} \right) \quad (15)$$

An important observation is that strategic complementarity $\alpha_{j,k,t}^{\gamma=0} \equiv (1 - \eta^{-1})m_{j,k,t}$ *increases* in the firm's own market share and *decreases* in the total market share of their competitors. This might seem unintuitive at first glance: after all, why should a firm's price be *more* sensitive to the prices of their competitors when those competitors hold *lower* market share? This becomes more puzzling in an extreme case when a single firm holds almost all the market with its market share approaching 1. The expression above implies that such a firm has the maximum strategic complementarity of $1 - \eta^{-1}$. But how can that be? Shouldn't a firm that holds almost all the market simply disregard their competitors and act as a monopoly?

The answer relies on the structure of demand implied by CES preferences: these preferences are such that the marginal consumer shifts away a higher share of her demand with respect to a one percent change in the prices of a firm's competitors when that firm holds higher market share. Thus, while a completely monopolistic firm enjoys the sheer lack of competition, the mere existence of small competitors shatters the autonomy of a firm in responding to their marginal costs, especially at higher levels of market share. Therefore, while a monopolistic firm with CES demand would charge a constant markup over its marginal cost, an *almost* monopolistic firm chooses to match the average price change of their competitors with weight $1 - \eta^{-1}$.

In the other extreme, when $m_{j,k,t}$ becomes small, the expression above becomes arbitrarily small and strategic complementarity disappears. This is not consistent with my findings in the empirical section of the paper where firms with a large number of competitors, and hence potentially lower market share, still report high levels of strategic complementarity.

This suggests that the sensitivity of markups is not the sole determinant of complementarities across firms and other forces might be at work. I capture this in the model by introducing decreasing returns to scale in labor, which is a standard approach in monetary models, especially in the absence of oligopolistic competition.

Nonetheless, the intuition outlined for the case of $\gamma = 0$ carries on to the case when $\gamma > 0$. With some tedious algebra in differentiating the best response of the firm with decreasing returns to scale, we can obtain the expression for strategic complementarity as

$$\alpha_{j,k,t}^{\gamma>0} = (1 - \eta^{-1})m_{j,k,t} + \left(1 - (1 - \eta^{-1})m_{j,k,t}\right) \left(1 - \frac{1 + \gamma}{1 + \gamma\eta(1 - (1 - \eta^{-1})m_{j,k,t})^2}\right) \quad (16)$$

This exposition of the strategic complementarity shows that at high levels of market share, the strategic complementarity is mainly driven by the sensitivity of the markup and gets closer to the strategic complementarity for the case of $\gamma = 0$ as $m_{j,k,t} \rightarrow 1$. However, now when $m_{j,k,t}$ becomes small, strategic complementarity remains positive and converges to $\frac{\gamma(\eta-1)}{1+\gamma\eta}$ as a firm's market share goes to zero.

4.3 Solution Method and Incentives in Information Acquisition

An Approximate Problem. I use a second-order approximation to the firms' problem to solve the model. The justification for this assumption stems from the issue that the problem of the firms, as stated in the previous section, has the joint distribution of prices and fundamental shocks its state variable. This is a known dimensionality curse in decision-making models of rational inattention that is exacerbated in the case of this model by the game-theoretic nature of firms' decisions as the solution requires solving for an additional fixed point across best response distributions.

Second-order approximations are a common remedy to this problem in the literature.²⁸ It is a well-known property of rational inattention models that when payoffs are quadratic and priors are Gaussian, optimal distributions are also Gaussian.²⁹ Since Gaussian distributions are characterized by their first and second moments, this approximation reduces the dimensionality of the problem to the squared dimension of the space from which the Gaussian distribution is drawn.³⁰

I derive this second-order approximation around the full-information equilibrium of this economy, which can be thought of as the case where information acquisition is free. Since

²⁸See, e.g., Maćkowiak and Wiederholt (2009, 2015); Paciello and Wiederholt (2014); Mackowiak et al. (2018).

²⁹See, e.g. Sims (2003); Maćkowiak and Wiederholt (2009); Afrouzi and Yang (2019) for proofs of this result in different environments.

³⁰See Afrouzi and Yang (2019) for a detailed discussion.

prices are flexible, classical dichotomy holds in the full-information economy and output is independent of monetary policy shocks. Thus, the symmetric equilibrium of this economy under full information is simply characterized by

$$P_{j,k,t}^{\text{full}} \propto Q_t, \forall j \in J, k \in K_j, t \geq 0 \quad (17)$$

Given this approximation, in Appendix D, I derive the implied approximate problem of the firm as

$$\max_{\{\kappa_{j,k,t}, S_{j,k,t}, p_{j,k,t}(S_{j,k}^t)\}_{t \geq 0}} -\mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t \left(\underbrace{B_j(p_{j,k,t}(S_{j,k}^t) - p_{j,k,t}^*)^2}_{\text{loss from mispricing}} + \underbrace{\omega \kappa_{j,k,t}}_{\text{cost of capacity}} |S_{j,k}^{-1}| \right) \right] \quad (18)$$

$$\begin{aligned} \text{s.t. } p_{j,k,t}^* &\equiv (1 - \alpha_j)q_t - \alpha_j p_{j,-k,t}(S_{j,-k,t}) \\ \mathcal{I} \left(S_{j,k,t}, (q_\tau, p_{l,m,\tau}(S_{l,m}^\tau))_{0 \leq \tau \leq t}^{(l,m) \neq (j,k)} \right) &\leq \kappa_{j,k,t} \\ S_{j,k}^t &= S_{j,k}^{t-1} \cup S_{j,k,t}, \quad S_{j,k}^{-1} \text{ given.} \end{aligned} \quad (19)$$

where small letters denote the log of their corresponding variables. Moreover, $B_j \equiv \frac{\eta + \gamma(\eta - (\eta - 1)K_j^{-1})^2}{1 + \gamma}$ is the curvature of firms' profit functions in sector j around their optimal price.

Information Acquisition Incentives. This approximate problem brings out the trade-offs that a firm faces in information acquisition. First, it formulates the profits of a firm as a function of two negative terms. The first term is the firm's losses from mispricing, which captures the fact that under imperfect information, the firm might choose a price that is not optimal under full information. This incorporates all the benefits of information acquisition: more information allows the firm to choose a price that is closer to the optimum, on average. The second term is the cost of producing information processing capacity, which increases with more information acquisition. The cost-benefit analysis between these two determines, first, the optimal capacity that the firm chooses for its information processing and, second, the signals that provide the firm with the best possible information given that capacity.

An important observation is that there is heterogeneity in the relative importance of losses from mispricing and the cost of producing information capacity. Firms with more competitors have more concave profit functions, which motivates them to produce more capacity, even though the cost of producing capacity is the same across all firms. This creates a level effect in information acquisition that was absent in the static model where capacity

was assumed to be constant. I will revisit this effect in more detail in Section 5.³¹

Furthermore, the problem formulates the losses from mispricing in terms of a quadratic distance from an “ideal” price ($p_{j,k,t}^*$) that is a weighted average between the log of nominal demand and the average price of the firm’s competitors. In the expression for this ideal price, the magnitude of the weight on others’ average price determines the strength of strategic incentives in information acquisition and is given by the degree of strategic complementarity, i.e. Equation (16), when market shares within the sector are symmetric and equal to $1/K_j$. The symmetric market shares assumption simplifies the solution to the model by allowing us to solve for the equilibrium among symmetric strategies. I discuss this assumption further in Section 6.

Finally, in addition to the strategic incentives discussed in the static model, the **evolution of the information set** over time indicates that how firms’ information set becomes the source of a new dynamic trade-off. At each period, firms understand that the signals they choose to observe will not only inform them about their contemporaneous ideal price, but also about its future values, as long as the process is autocorrelated.³² These dynamic incentives have two potential effects on information acquisition. (1) They affect the level of capacity production as a function of the volatility and persistence of “ideal” prices: a more patient firm that faces a persistent process assigns a larger continuation value to the knowledge that contemporaneous signals generate about the future values of said process. (2) Dynamic incentives also affect the allocation of attention between monetary policy shocks and others’ mistakes given their relative volatility and persistence: since mistakes are transitory but monetary policy shocks affect the level of prices forever, a more patient firm allocates more of its attention to monetary policy shocks.

Solving Firms’ Problem. Given a strategy profile for others, and the joint stochastic process that it implies for the vector of prices and the nominal demand, the solution to the problem of a firm is a joint stochastic process for the vector of prices of all firms in the economy along with the monetary policy shocks that solves the firm’s problem. The symmetric equilibrium is then a strategy profile from which no one has an incentive to deviate.

Appendix G thoroughly discusses my approach for solving for this joint stochastic process. Here, I mainly discuss the outline of the algorithm. I start by guessing a joint stochastic process for the prices in every sector. Given that the firms’ problems within sectors are sym-

³¹There is evidence that supports this level effect. Coibion et al. (2018) document that firms with a higher slope in their profit function around their optimal price have more accurate expectations about inflation.

³²See, e.g., Steiner et al. (2017); Mackowiak et al. (2018); Miao et al. (2020); Afrouzi and Yang (2019) for an extensive discussion of dynamic incentives of a rationally inattentive agent. Moreover, the exposition of dynamic rational inattention problems varies across different applications. For the formulation that is closest to this paper, see Afrouzi and Yang (2019).

metric, I then derive the implied strategy for the competitors of a representative firm in that sector given the initial guess. This strategy then implies a stochastic process for the “ideal” price of the firm specified in Equation (19). In particular, strategic inattention implies that this process depends on the history of monetary policy shocks as well as the history of non-fundamental shocks (mistakes) in the prices of a firm’s competitors. The state-space of the shocks that a firm desires to learn about is then given by:

$$p_{j,-k,t}(S_{j,-k}^t) = \underbrace{p_{j,-k,t}(S_{j,-k,t})|q}_{\text{projection on realizations of all } q_{t-\tau}'\text{s}} + \underbrace{v_{j,-k,t}}_{\text{orthogonal to all realizations of } q_{t-\tau}'\text{s}} \quad (20)$$

This equation also highlights a major difference between this model and a model in which there is a continuum of firms in which mistakes are orthogonal. In the latter, there is no non-fundamental volatility and $v_{j,-k,t} = 0$. The presence of these non-fundamental shocks motivates firms to learn about them and hence they need to be included in the state space of the shocks that firms track, where both their volatility and auto-correlations are endogenously determined.

Given the process for $p_{j,k,t}^*$ and its state-space representation, the problem of the firm then becomes a single agent dynamic rational inattention problem in a linear quadratic Gaussian setup that I solve using the methods developed in [Afrouzi and Yang \(2019\)](#). The solution to this problem characterizes the joint stochastic process of the firm’s price with the prices of its competitors and the monetary policy shocks, which constitutes the new guess for the equilibrium joint stochastic process. The solution method in [Appendix G](#) outlines and uses this algorithm to solve for the fixed point of this mapping among stochastic processes.

4.4 A Special Case with a Closed-Form Phillips Curve

In general, the equilibrium signal structure of firms does not admit a closed-form representation. However, we can go further in characterizing the representation of optimal signals when firms are completely myopic in their information acquisition ($\beta = 0$) which is useful for intuition.

Proposition 3. *Given a strategy profile for all other firms in the economy, every firm prefers to see only one signal at any given time. Moreover, if $\beta = 0$, the optimal signal of firm j, k at time t is*

$$S_{j,k,t} = (1 - \alpha_j)q_t + \alpha_j p_{j,-k,t}(S_{j,-k}^t) + e_{j,k,t}$$

The expression for the optimal signals in this case shows how firms incorporate the mistakes of their competitors into their information sets. In particular, by plugging in the

decomposition in Equation (20) we can re-write the optimal signal of the firm as

$$S_{j,k,t} = \underbrace{(1 - \alpha_j)q_t + \alpha_j p_{j,-k,t}(S_{j,-k,t})|_q}_{\text{predictive of } q_{t-\tau}'\text{s}} + \alpha_j v_{j,-k,t} + e_{j,k,t}.$$

predictive of sector price changes

This decomposition of the signal illustrates the main departure of this paper from models that assume a measure of firms. Since $\text{var}(v_{j,-k,t}) \neq 0$, the signal of a firm covaries more with the price changes of its competitors than with the fundamentals of the economy. When there is a measure of firms, however, the term $\alpha_j v_{j,-k,t}$ disappears and these two covariances converge to one another. Intuitively, this implies that when α_j is larger, firms in that sector are more informed about their own sector prices than the fundamentals of the economy.

We can go further in our special case and derive a closed-form expression for the Phillips curve of this economy for the case when there is no heterogeneity in the number of competitors across sectors. These assumptions are only made for the illustration of the Phillips curve and I will revert to the general case later in the calibrated model.

Proposition 4. *Suppose $\beta = 0$ and $K_j = K, \forall j \in J$ for some $K \in \mathbb{N}$. Then, $\alpha_j = \alpha, \forall j \in J$ and in the stationary equilibrium $\kappa_{j,k,t} = \kappa > 0, \forall j \in J, k \in K$. Moreover, the Phillips curve of this economy is*

$$\pi_t = (1 - \alpha) \overline{\mathbb{E}_{t-1}^{j,k}[\Delta q_t]} + \alpha \overline{\mathbb{E}_{t-1}^{j,k}[\pi_{j,-k,t}]} + (1 - \alpha)(2^{2\kappa} - 1)y_t,$$

where $\overline{\mathbb{E}_{t-1}^{j,k}[\Delta q_t]}$ is the average expected growth of nominal demand at $t - 1$, which is the sum of inflation and output growth, $\Delta q_t = \pi_t + \Delta y_t$, $\overline{\mathbb{E}_{t-1}^{j,k}[\pi_{j,-k,t}]}$ is the average expectation across firms of their competitors' price changes, and y_t is the output gap.

This Phillips curve crystallizes one of the main insights of this paper. In economies with large micro-level strategic complementarities, it is the firms' average expectation of their own competitors' price changes that drives aggregate inflation rather than their expectations of the growth in aggregate demand.

Moreover, Proposition 3 shows that with endogenous information acquisition, a larger α also implies that firms learn more about the prices of their competitors relative to the aggregate demand. Therefore, when α is large, not only is inflation driven more by firms' expectations of their own competitors' price changes but also firms' expectations are formed under information structures that are more informative about those prices.

Additionally, the slope of the Phillips curve shows how these strategic complementarities, as well as the capacity for processing information, affect monetary non-neutrality in this economy. The higher capacity of processing information makes the Phillips curve steeper, such that in the limit when $\kappa \rightarrow \infty$ (which arises endogenously when $\omega \rightarrow 0$), the

Phillips curve is vertical. In contrast, higher strategic complementarity makes the Phillips curve flatter since firms' higher-order beliefs become more important in their pricing decisions (Woodford, 2003a).

4.5 Calibration

The model is calibrated to the firm-level survey data from New Zealand at quarterly frequency, with a discount factor $\beta = 0.96^{1/4}$. A calibration to the US data would be ideal; however, one needs microdata on firms' expectations about inflation to calibrate the cost of attention in the US as well as data on how many competitors firm *directly* face to calibrate the distribution of the number of competitors, none of which are available for the US to my knowledge. For all the other parameters, however, the calibrated values to New Zealand data are also relatively in line with standard calibrations of these parameters to the US data. I discuss these case by case in the remainder of this section. Table (4) presents a summary of the calibrated values for all the parameters.

Table 4: Calibration Summary

<i>Parameter</i>	<i>Description</i>	<i>Value</i>	<i>Moment Matched</i>
\mathcal{K}	Distribution of K	$\sim \hat{\mathcal{K}}$	Empirical distribution (Fig. 1)
η	Elasticity of substitution	12	Elasticity of markups to $1/(1 - K_j^{-1})$
$1/(1 + \gamma)$	Curvature of production	0.526	Average strategic complementarity
ρ	Persistence of Δq	0.7	Persistence of NGDP growth in NZ
σ_u	Std. Dev. of shock to Δq	0.027	Std. Dev. of NGDP growth in NZ
ω	Cost of attention	0.326	Weight on prior in inflation forecasts

Notes: the table reports the calibrated values of the parameters for the dynamic model.

Distribution of the number of firms within oligopolies. I match the distribution of K_j in the model, denoted by \mathcal{K} , to the empirical distribution of the number of competitors that firms report in the data (Figure 1). As far as I know, there is no data available on how many competitors firms directly face in their market for the US.³³

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,

³³It is important to note that the value of K in this model corresponds to direct competitors of a firm that are only a small subset of all the firms that operate in a single SIC classification. Market segmentation, such as spatial constraints for consumers, make the number of firms within a SIC classification not suitable for calibrating this model. For instance, a coffee shop in Manhattan only competes with a small number of coffee shops that are geographically close to it, rather than all coffee shops in the US.

however, markups depend on the number of competitors and in the steady state are given by

$$\mu_j = 1 + \frac{1}{(\eta - 1)(1 - K_j^{-1})} \quad (21)$$

where K_j is the number of competitors in j . The survey measures the average markup by asking firms 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.”

Table 5: Calibration of η

	Average markup			
	(1)		(2)	
$1/(1 - K_j^{-1})$	0.107	(0.016)	0.123	(0.017)
Manufacturing			0.040	(0.007)
Professional and Financial Services			0.169	(0.007)
Trade			0.027	(0.007)
Constant	1.205	(0.018)	1.106	(0.020)
Observations	3152		3152	

Standard errors in parentheses

Notes: the table reports the result of regressing the average markups of firms on $1/(1 - K_j^{-1})$. The coefficient on this statistic is $1/(\eta - 1)$ in the model.

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 (21) and set $\eta = 12$ to match the coefficient on $\frac{1}{1 - K_j^{-1}}$ in the regression. Table (5) reports the result of this regression. This value is well in line with the values used in the literature for the US.

Curvature of production function. Given the empirical distribution of the number of firms, \mathcal{K} , and the elasticity of substitution, $\eta = 12$, I set $\gamma = 0.9$ to match the average degree of strategic complementarity $\bar{\alpha} = 0.8$ from Table (1). Given this value, the elasticity of output to labor in the model is 0.52. 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 et al., 2019](#), where the targeted value for the U.S. is 0.518).³⁴

Persistence and variance of shocks to nominal demand. I calibrate $\rho = 0.7$ to match the persistence of the growth of nominal GDP in New Zealand for post 1991 data.³⁵ Calibrated values for the US are slightly below this value and are in the range of 0.6 ([Midrigan, 2011](#)). However, the model is not very sensitive to this parameter in this range. For robustness, I present results for an alternative value of ρ in Section 6.

Given the quarterly persistence, I then set $\sigma_u = 0.027$ to match the unconditional variance of quarterly nominal GDP growth.³⁶ 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.³⁷

Cost of information acquisition. My strategy here is to target the weight that firms put on their priors in their inflation forecasts, which is an indicator of the degree of information rigidity and is similar to the approach in [Wiederholt \(2015\)](#). In particular, the survey follows a subset of firms across different waves and asks them about their yearly inflation forecasts (inflation 12 months ahead) and yearly inflation nowcasts (inflation in the previous 12 months). These horizons collapse for the first and the fourth waves of the survey that were conducted 12 months apart from one another (2013:Q4 to 2014:Q4). Thus, for the subset of firms that are present in both of these surveys, we observe their ex-ante and ex-post beliefs about inflation in that year. Kalman filtering implies that for firm i this revision should be given by

$$\mathbb{E}_{i,t}[\pi_t] = \mathbb{E}_{i,t-4}[\pi_t] + \lambda_i(s_{i,t} - \mathbb{E}_{i,t-4}[s_{i,t}]) \quad (22)$$

³⁴Although 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.

³⁵This 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 $\rho = 0.25^{1/4}$. 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.

³⁶The unconditional variance is given by $\frac{\sigma_u^2}{1-\rho^2}$ which is 0.0014 in the data.

³⁷This 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.

where t is in quarters, λ_i is the Kalman gain of the firm and $s_{i,t}$ is the signal(s) that the firm has observed within the year. The smaller the λ_i the more weight firms put on their priors and information rigidity is larger. Hence, under full information, when $\lambda_i = 1$, all firms should report the realized inflation and their priors should be irrelevant to their ex post nowcasts.

I follow the methodology of [Coibion and Gorodnichenko \(2015\)](#) to measure the degree of information rigidity in forecasts of aggregate inflation from the data. This methodology builds on the assumption that $s_{i,t} = \pi_t + noise$ and reduces the above updating equation to:

$$\mathbb{E}_{i,t}[\pi_t] = (1 - \lambda_i^*)\mathbb{E}_{i,t-4}[\pi_t] + \lambda_i^*\pi_t + error \quad (23)$$

Given this Equation, one can estimate the weight that firms put on their priors and back out the Kalman gain. However, note that the assumption $s_{i,t} = \pi_t + noise$ is incorrect based on the model since signals are endogenous in the model. If the model is the true data generating process, the identified coefficient on the prior is no longer the true λ_i but covaries with it – no matter what the true signals are, it is still the case that the prior should be less important when signals are more informative.

Since we do not observe the true signals of the firms in the data, we cannot control for the true signals; however, one can run the same misspecified regression within data generated by the model under different values of ω and choose the value that generates the same coefficient as in the data. In particular, I run the following regression:

$$\mathbb{E}_{i,t}[\pi_t] = constant + \delta\mathbb{E}_{i,t-4}[\pi_t] + error \quad (24)$$

where δ is the coefficient of interest. Since the regression exploits cross-sectional variation, rather than time-series variation as in [Coibion and Gorodnichenko \(2015\)](#), the current value of inflation is absorbed by the constant. Column (1) in Table (6) reports the baseline estimates for this specification. One caveat with this estimate is that there might be inherent heterogeneity among firms in perceiving different long-run inflation rates ([Patton and Timmermann, 2010](#)) which might get picked up by their ex-ante forecasts. A question in the survey asks firms about this target and Column (2) controls for this value.³⁸

To find the value for ω that generates the same coefficient in the model, I simulate the

³⁸The exact question is “What annual percentage rate of change in overall prices do you think the Reserve Bank of New Zealand is trying to achieve?” While this does not necessarily have to comply with firms’ beliefs about long-run inflation, a follow-up question verifies that they do. When asked “Do you believe the Reserve Bank of New Zealand can achieve its target?” the overwhelming majority of firms (more than 90 percent) respond yes. See [Kumar et al. \(2015\)](#) for a detailed discussion.

model for a range of values of ω . Figure (3) shows that ω is identified as this regression coefficient is sensitive to the cost of attention and increases in ω within the model. The fact that this coefficient increase with ω suggests that a high weight on the prior is associated with a larger cost of attention. I choose $\omega = 0.326$ to match the coefficient in Column (2) of Table (6) which yields a more conservative calibration relative to the coefficient in Column (1) – a value of ω matched to Column (1) would imply a larger cost of attention and a larger degree of monetary non-neutrality.

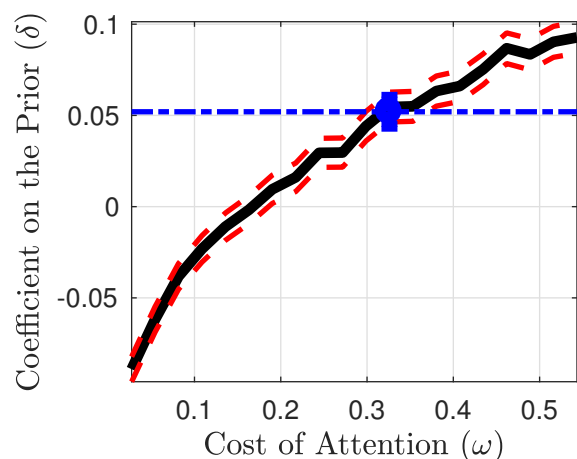


Figure 3: Sensitivity of δ to Cost of Attention (ω)
Notes: the black line shows the predicted value of δ from the regression specified in Equation (24) in model generated data as a function of ω . The blue dot shows the equivalent estimate in the New Zealand data from Table 6.

Table 6: Calibration of Cost of Attention (ω)

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

Standard errors in parentheses

Notes: the table reports the result of regressing firms' nowcast of yearly inflation on their forecast for the same horizon from 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 and gives a more conservative estimate.

To assess how this value for ω compares to other estimates of the degree of information rigidity in the literature, one can compare the implied Kalman gain from Equation (22) with the values that have been documented in the literature for professional forecasters. The average firm in this model has a Kalman gain of 0.51 which is above the estimated value for Professional Forecasters in the US by Coibion and Gorodnichenko (2015), who find an average Kalman gain of 0.45. Hence, the model implies that firms are more informed about their optimal prices than *professional forecasters are of aggregate inflation*. Nonetheless, firms exhibit large degrees of information rigidity in their inflation forecasts because inflation does not matter much for their decisions and their optimal signals are not as informative of inflation as they are of firms' optimal prices.

5 Results

5.1 Non-Targeted Moments: Subjective Uncertainty in the Model

The most crucial set of moments in the model that are not targeted in the calibration are given by the relationship between the subjective uncertainty of firms about aggregate inflation and the number of their competitors. As documented in the previous section, firms' uncertainty in the data is decreasing with the number of their competitors (Table 3). This relationship is not consistent with the benchmark models and is a unique property of this model that arises due to the interaction between rational inattention and oligopolistic competition in an environment that incorporates the heterogeneity in competition at the micro-level.

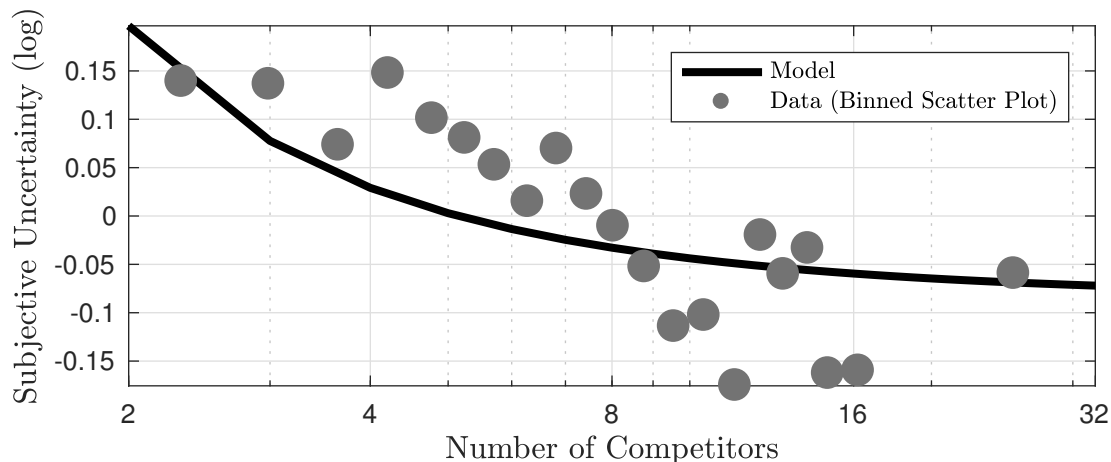


Figure 4: Subjective uncertainty about inflation: Model vs. Data.

Notes: the figure presents the fit of the model for the relationship between firms' (log) subjective uncertainty about inflation and the number of their competitors. The dots shows the binned scatter plot of log-subjective uncertainty about aggregate inflation against the number of competitors in the data. The black line depicts this relationship in the calibrated model. The average uncertainty is normalized to one in both the data and the model. This relationship was not targeted in the calibration of the model.

Figure (4) shows this relationship both in the model and the data (binned scatter plot).³⁹ The main observation is that the model matches the decline of the subjective uncertainty as a function of the number of competitors pretty well. It is important to note that heterogeneity in the number of competitors and endogenous information acquisition is key for this

³⁹I have normalized average uncertainty both in the data and in the model to 1.

relationship in the model: the former creates the differential incentives for information acquisition and the latter is essential for the endogenous variation in information acquisition.

Figure (5) shows the equilibrium level of firms' information acquisition and their implied Kalman gains as a function of the number of firms' competitors. More competitive firms (1) produce a higher capacity for processing information and (2) allocate more capacity towards aggregate shocks. As a result, more competitive firms have more accurate posteriors about aggregate variables, e.g. inflation.

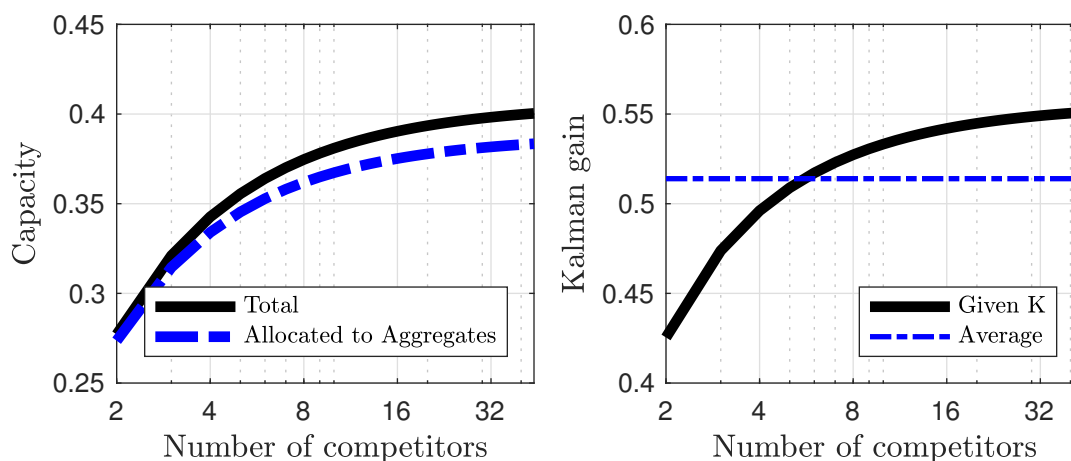


Figure 5: Information Capacity and Kalman Gains for Different Values of K .

Notes: the left panel shows the produced information processing capacity of a firm as a function of the number of competitors within its sector. The right panel shows the model implied true 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.

5.2 Implications for Monetary Non-Neutrality

The main driving force of my analysis so far has been the effect of a firm's number of competitors on their information acquisition incentives. In this section, I further this analysis by investigating how competition affects monetary non-neutrality and the propagation of monetary policy shocks to inflation and output. To do so, I will present two measures across different models.

(1) My first measure, following [Nakamura and Steinsson \(2010\)](#), is the variance of output and inflation to assess the degree of monetary non-neutrality. Since monetary policy shocks are the only shocks in my model, the natural level of output in the model is constant and thus any variation in output corresponds to a higher degree of monetary non-neutrality.

Similarly, a lower variance of inflation corresponds to a more muted response of inflation to monetary shocks. (2) My second measure is the cumulative half-life of output and inflation responses (time until the area under the impulse response reaches half of its full cumulative response) to monetary shocks to assess the persistence of the effect of monetary policy on these two variables.⁴⁰ The results of these comparisons are discussed in the remainder of this section and summarized in Tables (7) and (8).

Table 7: Output and Monetary Non-Neutrality Across Models

<i>Model</i>		<i>Variance</i>		<i>Persistence</i>	
		$\text{var}(Y) \times 10^3$	<i>amp. factor</i>	<i>half-life</i> ^{qtrs}	<i>amp. factor</i>
		(1)	(2)	(3)	(4)
Monopolistic Competition		1.53	1.00	3.38	1.00
Benchmark	$K \sim \hat{K}$	2.00	1.30	3.93	1.16
2-Competitors	$K = 2$	2.71	1.77	4.39	1.30
4-Competitors	$K = 4$	2.10	1.37	4.00	1.18
8-Competitors	$K = 8$	1.91	1.24	3.86	1.14
16-Competitors	$K = 16$	1.83	1.19	3.79	1.12
32-Competitors	$K = 32$	1.79	1.17	3.77	1.12
∞ -Competitors	$K \rightarrow \infty$	1.76	1.15	3.74	1.11

Notes: the table presents statistics for monetary non-neutrality across models with different number of competitors at the micro-level. $\text{Var}(Y)$ denotes the variance of output multiplied by 10^3 . *Half-life* denotes the length of the time that it takes for output to live half of its cumulative response in quarters. *Amp. factor* denotes the factor by which the relevant statistic is larger in the corresponding model relative to the model with monopolistic competition.

5.2.1 Comparison with the Monopolistic Competition Model

First, I compare the benchmark calibrated model to a model with monopolistic competition that has the same average degree of strategic complementarity in pricing. Equating the degree of strategic complementarity in pricing across the two models constitutes the right comparison in the sense that if I were to shut down strategic inattention in the benchmark model, the two models would yield the same impulse response functions. Moreover, it is also a desirable comparison since we know from a long line of previous papers that higher real rigidities (generated by strategic complementarity here) amplifies monetary non-neutrality (Ball and Romer, 1990) (in Section 5.3 below, we offer a more thorough

⁴⁰Usually, half-lives are measured as the time until a variable reaches half of its impact response. However, when responses are hump-shaped, as in this model, this can be misleading. To bypass this issue, I report the cumulative half-life.

Table 8: Inflation Across Models

Model	Variance		Persistence	
	$var(\pi) \times 10^4$ (1)	damp. factor (2)	half-life ^{qtrs} (3)	amp. factor (4)
Monopolistic Competition	9.27	1.00	3.94	1.00
Benchmark $K \sim \hat{K}$	8.73	0.94	4.34	1.10
2-Competitors $K = 2$	8.05	0.87	4.62	1.17
4-Competitors $K = 4$	8.62	0.93	4.38	1.11
8-Competitors $K = 8$	8.84	0.95	4.30	1.09
16-Competitors $K = 16$	8.94	0.96	4.26	1.08
32-Competitors $K = 32$	8.98	0.97	4.25	1.08
∞ -Competitors $K \rightarrow \infty$	9.03	0.97	4.23	1.07

Notes: the table presents statistics for inflation response across models with different number of competitors at the micro-level. $Var(\pi)$ denotes the variance of inflation multiplied by 10^4 . *Half-life* denotes the length of the time that it takes for inflation to live half of its cumulative response in quarters. *Damp. factor* (*amp. factor*) denotes the the factor by which the relevant statistic is smaller (larger) in the corresponding model relative to the model with monopolistic competition.

discussion on, and decomposition of, the confounding effect of strategic complementarities). The impulse response functions of output and inflation across these two models are presented in Figure (6).

The first two rows of Table (7) report how the behavior of output is different across these two models. Column (1) reports the variance of output across the two models.⁴¹ Column (2) reports the magnitude of amplification by normalizing the variance of output in the monopolistic competition to 1. Output is 30% more volatile under the benchmark model with strategic inattention. Column (3) reports the half-life of output response across the two models. While it takes 3.38 quarters for output to reach its half-life in the monopolistic competition model, this duration is 3.93 quarters in the benchmark model – a 16% increase as reported in Column (4).

The first two rows of Table (8) report how the behavior of inflation is different across these models. Inflation response is smaller and more persistent in the model with strategic inattention. Columns (1) and (2) show that inflation is 6% less volatile compared to the model with monopolistic competition. Furthermore, Column (3) shows that while it takes inflation 3.94 quarters to reach its cumulative half-life in the monopolistic competition model, this number is 4.34 quarters in the benchmark model – a 10% increase as reported in Column (4).

⁴¹Magnitudes are small since the variance of innovations to nominal GDP growth is small. The same is true for the US (Nakamura and Steinsson, 2010).

5.2.2 Counterfactual Distributions for Number of Competitors

So far I have presented results for a particular distribution of competition that is matched to the data in New Zealand, which was necessary for the calibration of the cost of attention. However, since the benchmark model targets this particular distribution, it masks the variation of monetary non-neutrality that comes from different degrees of granularity in the model. For this reason, I consider a second set of comparisons across homogeneous economies where every sector at the micro-level has $K \in \{2, 3, 4, \dots\}$ competitors. All other parameters are kept the same as in the benchmark model.

The results for output are listed in Table (7) and amplification factors are reported relative to the model with monopolistic competition. Monetary non-neutrality is larger and output response is more persistent in economies where the number of competitors is smaller. The duopoly model has the highest monetary non-neutrality, with an output volatility that is 77% larger than the monopolistic competition model, and a cumulative half-life that is a quarter longer.

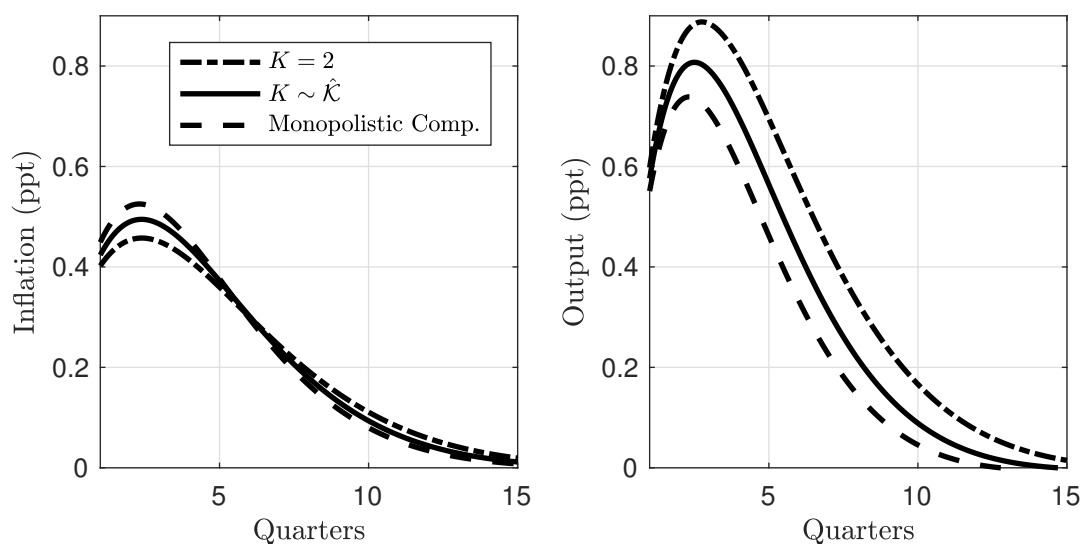


Figure 6: IRFs to a 1% Expansionary Shock

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 1). The dashed lines show the impulse responses in the model with monopolistic competition in all sectors. The dash-dotted lines show the impulse responses in the model where all sectors are composed of duopolies.

Table (8) reports the results for inflation. The smaller the number of competitors, the more muted is the response of inflation and the longer is its half-life. In the case of a duopoly within all sectors, the variance of inflation is 13% smaller than the model with monopolistic competition, and its half-life is 2 months (0.68 quarters) longer.

Figure (6) shows the impulse responses of output and inflation to a 1 percent unanticipated increase in the nominal aggregate demand for the benchmark model, the model with monopolistic competition and the economy where all sectors are duopolies. One important observation here is that the impulse responses of the model with $K \rightarrow \infty$ and monopolistic competition model are different, even though the two economies have an infinite number of competitors in each sector. The reason behind this is that the strategic complementarity in pricing (degree of real rigidity) depends on the number of competitors and is different across different levels of K . In particular, strategic complementarity is larger in more competitive economies and works in the *opposite* direction of strategic inattention. The next section explores these two effects in detail and aims at decomposing their contributions to monetary non-neutrality.

5.3 Decomposition: Strategic Inattention vs. Real Rigidities

Since the number of competitors affects both the degree of strategic complementarity and the amount of capacity produced by firms, the change in the degree of monetary non-neutrality across models with different numbers of competitors – as reported in the previous section – is the sum of two separate forces: (1) the real rigidity channel that alters monetary non-neutrality through changing the degree of strategic complementarity, and (2) the strategic inattention that alters the degree of non-neutrality through the amount of capacity produced and its allocation by firms.

Moreover, it is important to note that in the calibrated model, these two forces work in opposite directions: as discussed in Section 5.1 and shown in Figure (5), firms with a larger number of competitors produce higher capacity and allocate a larger amount of attention to learning about aggregates. Since firms with a larger allocated capacity towards aggregates learn monetary shocks more precisely and sooner, their prices move more swiftly in response to these shocks, and their output response is dampened as a result. Hence, monetary non-neutrality decreases with competition through the strategic inattention channel.

On the other hand, the degree of strategic complementarity increases with the number of competitors in the calibrated model, which follows from the expression of strategic complementarity in Equation (16). Figure (7) shows the degree of strategic complementarity in the calibrated model for different numbers of competitors.

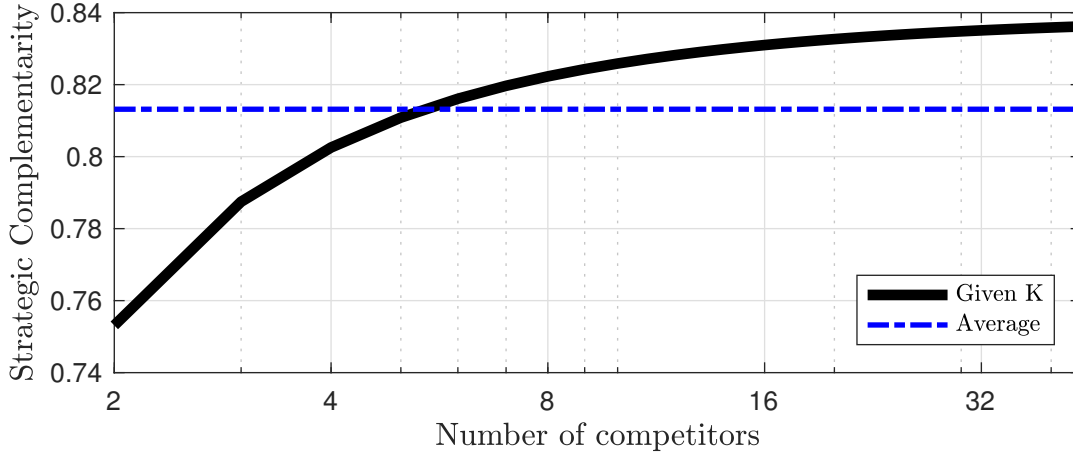


Figure 7: Strategic complementarity as a function of K .

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 higher degree of strategic complementarity. The dash-dotted line shows the average degree of strategic complementarity weighted by the empirical distribution of number of competitors in the New Zealand data.

Therefore, fixing the capacity of processing information, a larger number of competitors increases monetary non-neutrality in this model due to real rigidities by putting a larger weight on firms' higher-order beliefs (This is well-established in the literature of models with information rigidities. see, e.g. [Woodford, 2003a](#); [Nimark, 2008](#); [Mackowiak et al., 2018](#)). To verify this mechanism within the model, Figure (8) shows the IRFs of firms' higher-order beliefs to a one percent increase in nominal demand for three different values of K . For any given K , the responses of higher-order beliefs are smaller and more persistent. Therefore, when real rigidities are higher, firms put a larger weight on their higher-order beliefs and their average response also becomes smaller and more persistent, which amplifies monetary non-neutrality. Thus, monetary non-neutrality increases with the number of competitors through the real rigidity channel.

To decompose the effects of these two opposing forces, let us define $\alpha(K)$ to be the degree of strategic complementarity in a model where all sectors have K competitors and all the other parameters are fixed at their calibrated values. Moreover, let $\sigma_y^2(\alpha(K), K)$ denote the variance of output in the model where every sector has K competitors. The first argument captures the effect of the number of competitors on the weight that higher-order beliefs receive in the model (the real rigidity channel) and the second argument captures the effect of the number of competitors on the attention allocation of firms (strategic inattention channel). Then, we can decompose the difference in monetary non-neutrality of the

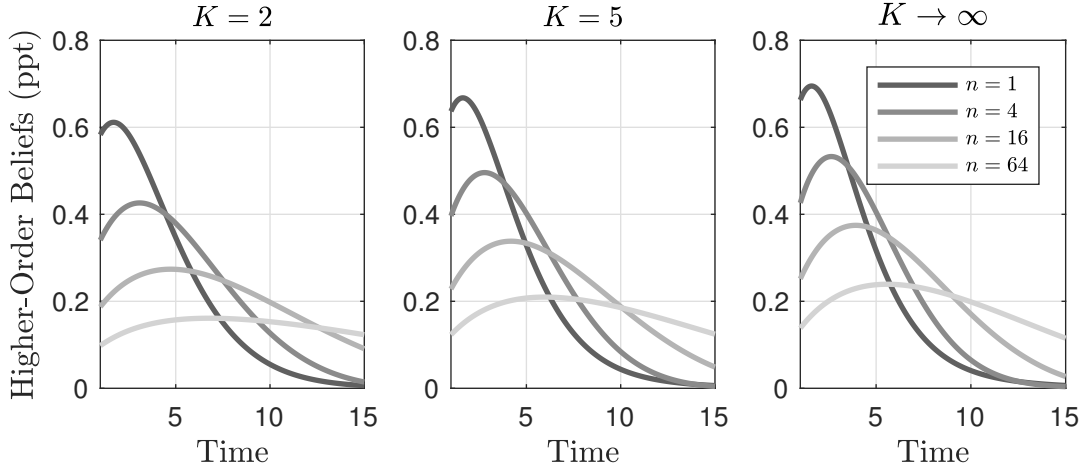


Figure 8: IRFs of Higher-Order Beliefs to a 1% Expansionary Shock

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 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.

two extreme models ($K = 2$ versus $K \rightarrow \infty$) as

$$\underbrace{\lim_{K \rightarrow \infty} \log \left(\frac{\sigma_y^2(\alpha(2), 2)}{\sigma_y^2(\alpha(K), K)} \right)}_{\text{total percentage change}} = \underbrace{\lim_{K \rightarrow \infty} \log \left(\frac{\sigma_y^2(\alpha(2), 2)}{\sigma_y^2(\alpha(2), K)} \right)}_{\text{percentage change due to strategic inattention}} + \underbrace{\lim_{K \rightarrow \infty} \log \left(\frac{\sigma_y^2(\alpha(2), K)}{\sigma_y^2(\alpha(K), K)} \right)}_{\text{percentage change due to real rigidities}} \quad (25)$$

Column (1) of Table (9) shows the results of this decomposition. Output variance is 43.3 percent larger with $K = 2$ relative to $K \rightarrow \infty$ (percentage difference here is calculated as the log-difference from Table (7)). Once decomposed to its two contributing factors, decreasing the number of competitors from $K \rightarrow \infty$ to $K = 2$ increases monetary non-neutrality by 85.6 percentage points due to the strategic inattention channel and decreases it by 42.3 percentage points through the real rigidity channel.

A similar decomposition can be done for inflation, whose variance is 11.5 percent smaller in the model with $K = 2$ relative to the model with $K \rightarrow \infty$. Column (2) of Table (9) shows that decreasing the number of competitors from $K \rightarrow \infty$ to $K = 2$ decreases the variance of inflation by 19.5 percentage points through the strategic inattention channel and increases it by 8 percentage points through the real rigidity channel.

Table 9: Decomposition: Strategic Inattention vs. Real Rigidities

	<i>Percentage change in variance of</i>	
	<i>output</i>	<i>inflation</i>
	(1)	(2)
Total Change (percent)	+43.3	-11.5
Due to Str. Inattention (ppt)	+85.6	-19.5
Due to Real Rigidities (ppt)	-42.3	+8.0

Notes: the table shows the decomposition of the effects of two opposing forces for the change in volatility of output (monetary non-neutrality) and inflation in the model with $K \rightarrow \infty$ versus the model with $K = 2$. Firms with higher number of competitors acquire more information about aggregate shocks, which decreases monetary non-neutrality and increases the volatility of inflation (strategic inattention channel). On the other hand, firms with larger number of competitors have larger degrees of strategic complementarity in pricing, which increases monetary non-neutrality and decreases volatility of inflation (real rigidity channel). For calibrated values of the parameters of the model, the strategic inattention channel dominates and monetary non-neutrality is 43.3% larger in the model with $K = 2$ than the model with $K \rightarrow \infty$.

6 Robustness and Alternative Assumptions

6.1 Lower persistence of nominal demand growth

In the calibration section, I argued that many of the parameter values that are calibrated to the New Zealand data are also consistent with the calibrations of those parameters for the US. One exception was the persistence of the nominal demand growth. While the value for this parameter is around 0.7 in New Zealand, its value in the US is around 0.5 to 0.6 (Nakamura and Steinsson, 2010; Midrigan, 2011; Mongey, 2018). To compare the results for this case, I redo the analysis for monetary non-neutrality for the case when $\rho = 0.5$. Figure (A.1a) in Appendix A shows the impulse response functions of inflation and output to a one percent expansion in the nominal aggregate demand under this assumption for the benchmark model, the model with 2 competitors in every sector and the monopolistic competition model. Moreover, Table (A.1a) in Appendix A reports the statistics for the volatility and persistence of output and inflation in this case.

The main takeaway is that while the amplification factors are slightly smaller than the case for $\rho = 0.7$, the results are fairly robust. For instance, output is 72% percent more volatile under the duopoly model relative to the model with monopolistic competition, which is only 5 percentage points smaller than the analogous number with $\rho = 0.7$.

6.2 Alternative discount factor

One of the main mechanisms in attention allocation within the model is the dynamic information acquisition of firms. Forward-looking firms internalize the continuation value of learning about different shocks and incorporate those incentives in their information acquisition. An important force here is that these dynamic incentives shift firms' attention towards more persistent processes because shocks to these processes are longer-lived (Afrouzi and Yang, 2019).

In the model, this mechanism dampens monetary non-neutrality. The reason is that monetary policy shocks are more persistent – due to the unit root in nominal demand – than the endogenous mistakes of firms, which are transitory because mistakes cannot persist forever. Thus, more forward-looking firms allocate more attention to the monetary policy shocks, which attenuates strategic inattention and reduces monetary non-neutrality.

Since I have calibrated the discount factor to a common value of $0.96^{0.25}$, which is very close to 1, the dynamic incentives are very strong. However, information is an intangible and, potentially, a non-tradable form of capital and the discount factor associated with this form of capital does not have to comply with the discount factor of households. As an alternative approach, I calibrate β and ω jointly by targeting the same moment in the data (the persistence of forecast errors as in the benchmark calibration) and redo the results for monetary non-neutrality. This joint calibration yields a value of 0.6 for β and a value of 0.217 for ω . An important observation is that the calibrated value for the cost of inattention, in this case, is *smaller* than the value in the benchmark calibration (0.217 versus 0.326) which means that firms face lower costs in acquiring information in this alternative calibration.

Figure (A.1b) in Appendix A shows the impulse responses of output and inflation to a one percent expansion in nominal aggregate demand for the benchmark model, the model with 2 competitors in every sector and the model with monopolistic competition. Moreover, Table (A.1b) in Appendix A reports the statistics for the volatility and persistence of output and inflation in this case.

The main takeaway is that even though the cost of attention is smaller in this calibration relative to the benchmark, monetary non-neutrality is larger. For instance, under this alternative calibration output volatility is 112% larger in the model with 2 competitors in every sector relative to the model with monopolistic competition, an amplification factor that is 35 percentage points larger than the analogous factor under the benchmark calibration (77%). The intuition behind this result is that in spite of the lower cost of attention, firms are myopic in information acquisition: they produce lower capacity than the benchmark calibration to begin with, and given that capacity, they allocate a higher share of that to the mistakes of their competitors since they are relatively more ignorant of the continuation

value of information.

6.3 Heterogeneity within sector market shares

In deriving the approximate problem in Section 4.3, I took a second-order approximation to the firms' problem around a symmetric steady state, in which all firms within the same sector had the same market share. This approximation makes solving the problem feasible by making problems of all firms within one sector symmetric. However, it ignores potential heterogeneity in market shares.

In particular, one important question is how does heterogeneity in market shares affect strategic inattention? For instance, consider a duopoly in which one firm holds almost all the market share where its rival has almost zero market share. Is it the case that the large firm ignores the mistakes of the small firm and allocates all of its attention to the aggregates? The short answer to this question is no. In fact, the opposite is true and the large firm pays more attention to the mistakes of the small firm, and the small firm pays almost full attention to the aggregates.

The reasoning behind this argument is that a large firm's optimal price with CES demand is more sensitive to the prices of its competitors at higher levels of market share, an evident observation from the expression of strategic complementarity in Equation (15). Therefore, firms with larger market shares face higher strategic complementarities and are more affected by the mistakes of smaller firms in their sector.

Furthermore, it is important to recognize the discontinuity that this implies for an oligopoly with a firm holding *almost* total market share and a monopoly: in an oligopoly where a single firm holds an arbitrarily large portion of the market, the mere existence of small rivals – who are ready to steal the market share of the larger firm if an opportunity presents itself – motivates the larger firm to pay a lot of attention to the beliefs of these rivals. This is different than the case of a monopoly where the firm does not have to worry about any off-equilibrium threat of small firms stealing its market share. A monopoly with a constant elasticity of demand always pays one hundred percent of its attention to shocks to its marginal cost.

While solving the quantitative model without the symmetric market share approximation is not feasible, we can investigate how the heterogeneity in market shares would matter by at least deriving the second order approximation in a simple model with heterogeneous market shares. Appendix H discusses a simple case with CES preferences and shows that, up to a second-order approximation, the strategic complementarity of any given firm is their *market share* in the steady state. Therefore, firms with higher market shares have higher strategic complementarities, and thus higher incentives to track others' mistakes rather than

the aggregate shocks.

Therefore, we expect the symmetric market share approximation to be a conservative estimate of the effect of oligopolistic competition on monetary non-neutrality. In non-symmetric cases, larger firms, who contribute more to the output of the economy, will pay less attention to monetary policy shocks and output is expected to be more volatile.

7 Concluding Remarks

In this paper, I develop a new quantitative model to address the link between oligopolistic competition and information acquisition. I find that the interaction of these two frictions creates an endogenous correlation between the accuracy of firms' beliefs and the number of their competitors. Oligopolistic firms find it optimal to pay direct attention to the beliefs of their competitors, an incentive that is stronger when the number of those competitors is smaller.

Moreover, I show that these endogenous strategic incentives in information acquisition have significant implications for the propagation of monetary policy shocks to output and inflation. In tracking their competitors' beliefs, firms ignore aggregate shocks and as a result respond to these shocks more slowly and sluggishly. Calibrating the model to firms-level survey data, I find that these strategic incentives increase monetary non-neutrality by up to 77% and increase the half-life of output to a monetary shock by 30%.

The results of this paper also provide valuable insights for policy. The link between competition and monetary non-neutrality suggests that the recently documented trends in competition (De Loecker et al., 2020; Autor et al., 2017) are also changing the landscape of monetary policy by affecting the propagation of these shocks to real and nominal variables.

Furthermore, the results in this paper have implications for policies that target expectations. In particular they provide a new perspective on why managing inflation expectations might be less effective than what a model with monopolistic competition would suggest. Managers of oligopolistic firms do not directly care about aggregate inflation and are mainly concerned with how their competitors change their prices in the face of a shock. As a result, they are more informed about their optimal prices than what their expectations of aggregate inflation would suggest.

The fact that aggregate inflation is not the primary concern of these firms implies that unanchored inflation expectations are not necessarily a problem for monetary policy. After all, the main objective of inflation targeting is to stabilize inflation, and a byproduct of such policies is that inflation will no longer be a concern for firms. Therefore, the fact that firms do not have to track it closely when it is low and stable is in itself a success for monetary

policy. However, this implies that managing expectations of aggregate inflation is neither an effective tool for controlling inflation nor necessarily a powerful instrument for policies such as forward guidance. These expectations are relatively unimportant for firms and do not have much impact on their pricing decisions.

Nevertheless, this result does not necessarily rule out policies that target expectations, but rather provides a new view on how those policies should be framed and *which* expectations they should target. An important takeaway from this paper is that for such a policy to be successful, it has to communicate the course of monetary policy to price-setters not in terms of how it will steer the overall prices but in terms of how it will affect their own industry prices. In other words, framing policy in terms of the aggregate variables will not gain as much attention and response from firms as it would if the news about the policy were to reach firms in terms of how their competitors would be affected. How policy can achieve these ends remains a question that deserves more investigation.

References

- Afrouzi, Hassan and Choongryul Yang**, “Dynamic Rational Inattention and the Phillips Curve,” 2019. Mimeo.
- Alvarez, Fernando E, Francesco Lippi, and Luigi Paciello**, “Optimal price setting with observation and menu costs,” *The Quarterly journal of economics*, 2011, 126 (4), 1909–1960.
- Alvarez, Fernando, Herve Le Bihan, and Francesco Lippi**, “The Real Effects of Monetary Shocks in Sticky Price Models: A Sufficient Statistic Approach,” *American Economic Review*, 2016, 106 (10), 2817–51.
- Angeletos, George-Marios and Alessandro Pavan**, “Efficient Use of Information and Social Value of Information,” *Econometrica*, 2007, 75 (4), pp. 1103–1142.
- **and Jennifer La’O**, “Incomplete Information, Higher-Order Beliefs and Price Inertia,” *Journal of Monetary Economics*, 2009, 56, S19–S37.
- Atkeson, Andrew and Ariel Burstein**, “Pricing-to-Market, Trade Costs, and International Relative Prices,” *American Economic Review*, 2008, 98 (5), 1998–2031.
- Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, John Van Reenen et al.**, “The Fall of the Labor Share and the Rise of Superstar Firms,” 2017. NBER Working Paper No. 23396.
- Baley, Isaac and Andrés Blanco**, “Firm Uncertainty Cycles and the Propagation of Nominal Shocks,” *American Economic Journal: Macroeconomics*, 2019, 11 (1), 276–337.
- Ball, Laurence and David Romer**, “Real Rigidities and the Non-Neutrality of Money,” *The Review of Economic Studies*, 1990, 57 (2), 183–203.

- Bilal, Adrien G, Niklas Engbom, Simon Mongey, and Giovanni L Violante**, "Firm and Worker Dynamics in a Frictional Labor Market," 2019. NBER Working Paper No. 26547.
- Bils, Mark and Peter J Klenow**, "Some Evidence on the Importance of Sticky Prices," *Journal of political economy*, 2004, 112 (5), 947–985.
- Bryan, Michael F, Brent Meyer, and Nicholas Parker**, "The Inflation Expectations of Firms: What Do They Look Like, Are They Accurate, and Do They Matter?," 2015. Working Paper No. 2014-27a.
- Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans**, "Monetary policy shocks: What have we learned and to what end?," *Handbook of macroeconomics*, 1999, 1, 65–148.
- Coibion, Olivier and Yuriy Gorodnichenko**, "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *The American Economic Review*, 2015, 105 (8), 2644–2678.
- , – , and **Saten Kumar**, "How Do Firms Form Their Expectations? New Survey Evidence," *American Economic Review*, 2018, 108 (9), 2671–2713.
- Cover, Thomas M and Joy A Thomas**, *Elements of Information Theory*, John Wiley & Sons, 2012.
- Denti, Tommaso**, "Games with Unrestricted Information Acquisition," 2018. Mimeo.
- Friedman, Milton**, "The Role of Monetary Policy," *The American Economic Review*, 1968, 58 (1), 1–17.
- Fulton, Chad**, "Mechanics of Static Quadratic Gaussian Rational Inattention Tracking Problems," Working Paper, Board of Governors of the Federal Reserve System 2018.
- Galí, Jordi**, *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*, Princeton University Press, 2015.
- Gertler, Mark and John Leahy**, "A Phillips Curve with an Ss Foundation," *Journal of political Economy*, 2008, 116 (3), 533–572.
- and **Peter Karadi**, "Monetary policy surprises, credit costs, and economic activity," *American Economic Journal: Macroeconomics*, 2015, 7 (1), 44–76.
- Golosov, Mikhail and Robert E Lucas Jr**, "Menu Costs and Phillips Curves," *Journal of Political Economy*, 2007, 115 (2), 171–199.
- Hellwig, Christian and Laura Veldkamp**, "Knowing What Others Know: Coordination Motives in Information Acquisition," *The Review of Economic Studies*, 2009, 76 (1), 223–251.
- and **Venky Venkateswaran**, "Setting the Right Prices for the Wrong Reasons," *Journal of Monetary Economics*, 2009, 56, S57–S77.

- Ilut, Cosmin and Rosen Valchev**, “Economic Agents As Imperfect Problem Solvers,” 2017. Manuscript.
- Jr, Robert E Lucas**, “Expectations and the Neutrality of Money,” *Journal of economic theory*, 1972, 4 (2), 103–124.
- Khaw, Mel Win and Oskar Zorrilla**, “Deeper Habits,” 2018. Manuscript.
- , **Luminita Stevens, and Michael Woodford**, “Discrete Adjustment to a Changing Environment: Experimental Evidence,” *Journal of Monetary Economics*, 2017, 91, 88–103.
- Kumar, Saten, Hassan Afrouzi, Olivier Coibion, and Yuriy Gorodnichenko**, “Inflation Targeting Does Not Anchor Inflation Expectations: Evidence from Firms in New Zealand,” *Brookings Papers on Economic Activity*, 2015, Fall, 151–208.
- Leeper, Eric M, Christopher A Sims, and Tao Zha**, “What Does Monetary Policy Do?,” *Brookings papers on economic activity*, 1996, 1996 (2), 1–78.
- Loecker, Jan De, Jan Eeckhout, and Gabriel Unger**, “The Rise of Market Power and the Macroeconomic Implications,” *The Quarterly Journal of Economics*, 2020.
- Luo, Yulei**, “Consumption Dynamics under Information Processing Constraints,” *Review of Economic Dynamics*, 2008, 11 (2), 366–385.
- , **Jun Nie, and Eric R Young**, “Robustness, Information Processing Constraints, and the Current Account in Small Open Economies,” *Journal of International Economics*, 2012, 88 (1), 104–120.
- Maćkowiak, Bartosz and Mirko Wiederholt**, “Optimal Sticky Prices under Rational Inattention,” *The American Economic Review*, 2009, 99 (3), 769–803.
- and – , “Business Cycle Dynamics under Rational Inattention,” *The Review of Economic Studies*, 2015, 82 (4), 1502–1532.
- Mackowiak, Bartosz, Filip Matejka, and Mirko Wiederholt**, “Dynamic Rational Inattention: Analytical Results,” *Journal of Economic Theory*, 2018, 176, 650 – 692.
- Mankiw, N. Gregory and Ricardo Reis**, “Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve,” *The Quarterly Journal of Economics*, 2002, 117 (4), pp. 1295–1328.
- Matějka, Filip**, “Rationally Inattentive Seller: Sales and Discrete Pricing,” *The Review of Economic Studies*, 2015, p. rdv049.
- and **Alisdair McKay**, “Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model,” *American Economic Review*, 2015, 105 (1), 272–98.
- Melosi, Leonardo**, “Signalling Effects of Monetary Policy,” *The Review of Economic Studies*, 2016, 84 (2), 853–884.

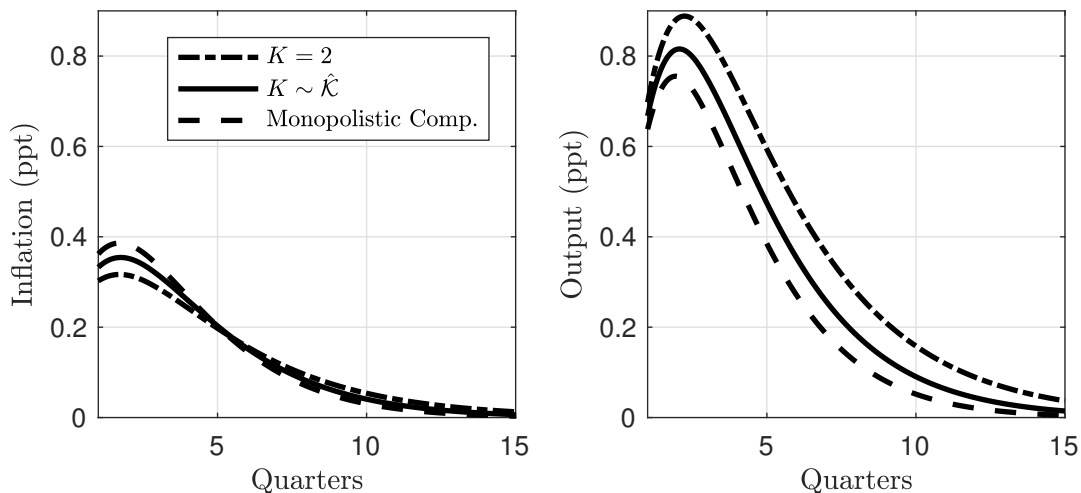
- Miao, Jianjun, Jieran Wu, and Eric Young**, “Multivariate Rational Inattention,” Technical Report, Working Paper, Boston University 2020.
- Midrigan, Virgiliu**, “Menu costs, multiproduct firms, and aggregate fluctuations,” *Econometrica*, 2011, 79 (4), 1139–1180.
- Mondria, Jordi and Thomas Wu**, “The Puzzling Evolution of the Home Bias, Information Processing and Financial Openness,” *Journal of Economic Dynamics and Control*, 2010, 34 (5), 875–896.
- Mongey, Simon**, “Market Structure and Monetary Non-Neutrality,” 2018. Mimeo.
- Morris, Stephen and Hyun Song Shin**, “Social Value of Public Information,” *The American Economic Review*, 2002, 92 (5), pp. 1521–1534.
- Nakamura, Emi and Jón Steinsson**, “Five Facts about Prices: A Reevaluation of Menu Cost Models,” *The Quarterly Journal of Economics*, 2008, 123 (4), 1415–1464.
- **and Jon Steinsson**, “Monetary Non-Neutrality in a Multisector Menu Cost Model,” *The Quarterly journal of economics*, 2010, 125 (3), 961–1013.
- **and Jón Steinsson**, “High-Frequency Identification of Monetary Non-Neutrality: the information effect,” *The Quarterly Journal of Economics*, 2018, 133 (3), 1283–1330.
- Nimark, Kristoffer**, “Dynamic Pricing and Imperfect Common Knowledge,” *Journal of monetary Economics*, 2008, 55 (2), 365–382.
- Paciello, Luigi and Mirko Wiederholt**, “Exogenous information, endogenous information, and optimal monetary policy,” *Review of Economic Studies*, 2014, 81 (1), 356–388.
- Pasten, Ernesto and Raphael Schoenle**, “Rational inattention, multi-product firms and the neutrality of money,” *Journal of Monetary Economics*, 2016, 80, 1–16.
- Patton, Andrew J and Allan Timmermann**, “Why Do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion,” *Journal of Monetary Economics*, 2010, 57 (7), 803–820.
- Peng, Lin and Wei Xiong**, “Investor Attention, Overconfidence and Category Learning,” *Journal of Financial Economics*, 2006, 80 (3), 563–602.
- Phelps, Edmund S**, “Phillips Curves, Expectations of Inflation and Optimal Unemployment over Time,” *Economica*, 1967, pp. 254–281.
- Romer, Christina D and David H Romer**, “A New Measure of Monetary Shocks: Derivation and Implications,” *American Economic Review*, 2004, 94 (4), 1055–1084.
- Rotemberg, Julio J and Michael Woodford**, “Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity,” *The Journal of Political Economy*, 1992, 100 (6).
- Shannon, Claude E**, “A Mathematical Theory of Communication,” *Bell System technical Journal*, 1948, 27, 379–423.

- Sims, Christopher A**, "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, 50 (3), 665–690.
- , "Rational Inattention: Beyond the Linear-Quadratic Case," *The American economic review*, 2006, 96 (2), 158–163.
- Steiner, Jakub, Colin Stewart, and Filip Matějka**, "Rational Inattention Dynamics: Inertia and Delay in Decision-Making," *Econometrica*, 2017, 85 (2), 521–553.
- Stevens, Luminita**, "Coarse Pricing Policies," *The Review of Economic Studies*, 2020, 87 (1), 420–453.
- Tutino, Antonella**, "Rationally Inattentive Consumption Choices," *Review of Economic Dynamics*, 2013, 16 (3), 421–439.
- Uhlig, Harald**, "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 2005, 52 (2), 381–419.
- Vavra, Joseph**, "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation," *The Quarterly Journal of Economics*, 2014, 129 (1), 215–258.
- Wiederholt, Mirko**, "Empirical Properties of Inflation Expectations and the Zero Lower Bound," 2015. Mimeo.
- Woodford, Michael**, "Imperfect Common Knowledge and the Effects of Monetary Policy," *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, 2003, p. 25.
- , *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, 2003.
- Yang, Choongryul**, "Rational Inattention, Menu Costs, and Multi-Product Firms: Evidence and Monetary Non-Neutrality," 2019. Manuscript.
- Zorn, Peter**, "Investment under Rational Inattention: Evidence from US Sectoral Data," 2016. Manuscript.

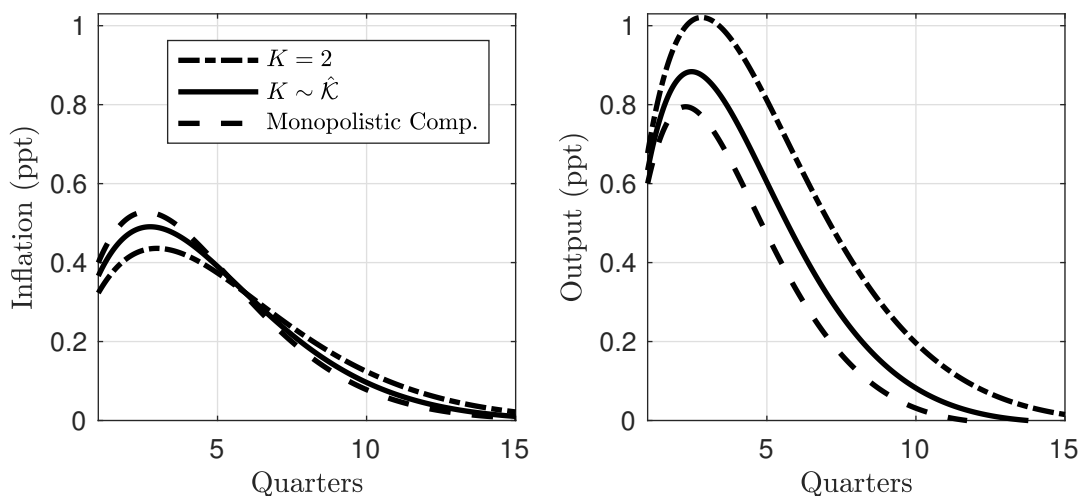
APPENDIX
(FOR ONLINE PUBLICATION)

A Additional Figures and Tables

Figure A.1: Robustness – Overall Effects of Oligopolistic Competition



(a) Alternative persistence for the growth of nominal aggregate demand ($\rho = 0.5$)



(b) Alternative discount rate for information ($\beta = 0.6$)

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 with alternative calibration of $\rho = 0.5$. 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 1). The dashed lines show the impulse responses in the model with monopolistic competition in all sectors. The dash-dotted lines show the impulse responses in the model where all sectors are composed of duopolies.

Table A.1: Robustness – Output and Inflation Across Models

(a) Alternative persistence for the growth of nominal aggregate demand ($\rho = 0.5$)

Model	Output				Inflation			
	Variance		Persistence		Variance		Persistence	
	$var(Y) \times 10^3$ (1)	<i>amp.</i> factor (2)	<i>half-life</i> q_{trs} (3)	<i>amp.</i> factor (4)	$var(\pi) \times 10^4$ (5)	<i>damp.</i> factor (6)	<i>half-life</i> q_{trs} (7)	<i>amp.</i> factor (8)
Monopolistic Competition	1.44	1.00	3.21	1.00	3.97	1.00	3.30	1.00
Benchmark $K \sim \hat{K}$	1.85	1.28	3.79	1.18	3.59	0.90	3.79	1.15
2-Competitors $K = 2$	2.48	1.72	4.29	1.34	3.13	0.79	4.19	1.27
4-Competitors $K = 4$	1.94	1.34	3.86	1.20	3.51	0.88	3.85	1.17
8-Competitors $K = 8$	1.76	1.22	3.71	1.16	3.66	0.92	3.74	1.13
16-Competitors $K = 16$	1.69	1.17	3.65	1.14	3.73	0.94	3.69	1.12
32-Competitors $K = 32$	1.66	1.15	3.62	1.13	3.76	0.95	3.67	1.11
∞ -Competitors $K \rightarrow \infty$	1.63	1.13	3.59	1.12	3.79	0.96	3.65	1.11

(b) Alternative discount rate for information ($\beta=0.6$)

Model	Output				Inflation			
	Variance		Persistence		Variance		Persistence	
	$var(Y) \times 10^3$ (1)	<i>amp.</i> factor (2)	<i>half-life</i> q_{trs} (3)	<i>amp.</i> factor (4)	$var(\pi) \times 10^4$ (5)	<i>damp.</i> factor (6)	<i>half-life</i> q_{trs} (7)	<i>amp.</i> factor (8)
Monopolistic Competition	1.71	1.00	3.25	1.00	9.31	1.00	3.99	1.00
Benchmark $K \sim \hat{K}$	2.34	1.37	3.84	1.18	8.65	0.93	4.44	1.11
2-Competitors $K = 2$	3.62	2.12	4.41	1.36	7.67	0.82	4.90	1.23
4-Competitors $K = 4$	2.51	1.47	3.91	1.20	8.50	0.91	4.51	1.13
8-Competitors $K = 8$	2.18	1.28	3.73	1.15	8.81	0.95	4.38	1.10
16-Competitors $K = 16$	2.05	1.20	3.66	1.13	8.95	0.96	4.32	1.08
32-Competitors $K = 32$	1.99	1.17	3.62	1.11	9.01	0.97	4.30	1.08
∞ -Competitors $K \rightarrow \infty$	1.94	1.13	3.59	1.10	9.07	0.97	4.27	1.07

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 ($\rho = 0.5$). Panel (b) presents results for an alternative calibration of discount rate for information ($\beta = 0.6$). $Var(\cdot)$ denotes the variance of output/inflation. *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.

B Proofs for the Static Model

This section formalizes the static game in Section 2. The Appendix is organized as follows. I start by specifying the Shannon mutual information function in Subsection B.1. Subsection B.2 defines the concept of *richness* for a set of available information, and characterizes such a set. The main idea behind having a rich set of available information is to endow firms with the freedom of choosing their ideal signals given their capacity. Following this, Subsection B.3 proves the optimality of linear pricing strategies given Gaussian signals, and Subsection B.4 proves that when the set of available signals is rich all firms prefer to see a *single* signal. Subsection B.5 shows that any equilibrium has an equivalent in terms of the joint distribution it implies for prices among the strategies in which all firms observe a single signal, and derives the conditions that such signals should satisfy. Subsection B.6 shows that the equilibrium is unique given this equivalence relationship. Subsection B.7 derives an intuitive reinterpretation of a firm's attention problem that is discussed in Section 2. Subsection B.8 contains the proofs of propositions and corollaries for the static model.

B.1 Shannon's Mutual Information

In information theory a mutual information function is a function that measures the amount of information that two random variables reveal about one another. In this paper following the rational inattention literature, I use Shannon's mutual information function for the attention constraint of the firms, which is defined as the reduction in entropy that the firm experiences given its signal.⁴² In case of Gaussian variables, this function takes a simple and intuitive form. Let $(X, Y) \sim \mathcal{N}(\mu, \begin{bmatrix} \Sigma_X & \Sigma_{X,Y} \\ \Sigma_{Y,X} & \Sigma_Y \end{bmatrix})$. Then, the mutual information between X and Y is given by $\mathcal{I}(X; Y) = \frac{1}{2} \log_2\left(\frac{\det(\Sigma_X)}{\det(\Sigma_{X|Y})}\right)$, where $\Sigma_{X|Y} = \Sigma_X - \Sigma_{X,Y}\Sigma_Y^{-1}\Sigma_{Y,X}$ is the variance of X conditional on Y . Intuitively, the mutual information is bigger if the Y reveals more information about X , leading to a smaller $\det(\Sigma_{X|Y})$. In the other extreme case where $X \perp Y$, then $\Sigma_{X|Y} = \Sigma_X$ and $\mathcal{I}(X; Y) = 0$, meaning that if 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 will use for proving the optimality of single signals is the *data processing inequality*. The following lemma proves a weak version of this inequality for completeness.

⁴²In his seminal paper Shannon (1948) showed that under certain axioms there is a unique entropy function.

Lemma B.1. Let X, Y and Z be three random variables such that $X \perp Z|Y$.⁴³ Then $\mathcal{I}(X; Y) \geq \mathcal{I}(X; Z)$.

Proof. By the chain rule for mutual information⁴⁴ $\mathcal{I}(X; (Y, Z)) = \mathcal{I}(X; Y) + \mathcal{I}(X; Z|Y) = \mathcal{I}(X; Z) + \mathcal{I}(X; Y|Z)$. Notice that since $X \perp Z|Y$, then $\mathcal{I}(X; Z|Y) = 0$. Thus, $\mathcal{I}(X; Y) = \mathcal{I}(X; Z) + \underbrace{\mathcal{I}(X; Y|Z)}_{\geq 0} \geq \mathcal{I}(X; Z)$. \square

B.2 A Rich Set of Available Information

Definition. Let \mathcal{S} be a set of Gaussian signals. We say \mathcal{S} is *rich* if for any mean-zero possibly multivariate Gaussian distribution G , there is a vector of signals in \mathcal{S} that are distributed according to G .

To specify a rich information structure, suppose in addition to $q \sim \mathcal{N}(0, 1)$ there are countably infinite independent Gaussian noises in the economy, meaning that there is a set $\mathcal{B} \equiv \{q, e_1, e_2, \dots\}$ such that $\forall i \in \mathbb{N}, e_i \sim \mathcal{N}(0, 1), e_i \perp q$ and $\forall \{i, j\} \subset \mathbb{N}, j \neq i, e_j \perp e_i$. Let \mathcal{S} be the set of all finite linear combinations of the elements of \mathcal{B} with coefficients in \mathbb{R} : $\mathcal{S} = \{a_0 q + \sum_{i=1}^N a_i e_{\sigma(i)}, N \in \mathbb{N}, (a_i)_{i=0}^N \subset \mathbb{R}^{N+1}, (\sigma(i))_{i=1}^N \subset \mathbb{N}\}$. We let \mathcal{S} denote the set of all available signals in the economy.

Lemma B.2. \mathcal{S} is rich.

Proof. Suppose G is a mean-zero Gaussian distribution. Thus, $G = \mathcal{N}(0, \Sigma)$, where $\Sigma \in \mathbb{R}^{N \times N}$ is a positive semi-definite matrix for some $N \in \mathbb{N}$. Since Σ is positive semi-definite, by Spectral theorem there exists $A \in \mathbb{R}^{N \times N}$ such that $\Sigma = A' \times A$. Choose any N elements of \mathcal{B} , and let \mathbf{e} be the vector of those elements. Then $\mathbf{e} \sim \mathcal{N}(0, \mathbf{I}_{N \times N})$ where $\mathbf{I}_{N \times N}$ is the N dimensional identity matrix. By definition of \mathcal{S} , $S \equiv A' \mathbf{e} \in \mathcal{S}$. Now notice that $\mathbb{E}[S] = 0$, $var(S) = A' var(\mathbf{e}) A = \Sigma$. Hence, $S \sim \mathcal{N}(0, \Sigma) = G$. \square

Definition. For a vector of non-zero Gaussian signals $S \sim \mathcal{N}(0, \Sigma)$, we say elements of S are *distinct* if Σ is invertible. In other words, elements of S are distinct if no two signals in S are perfectly correlated.

Corollary B.1. Let S be an N -dimensional vector of non-zero distinct signals whose elements are in \mathcal{S} . Let $G = \mathcal{N}(0, \Sigma)$ be the distribution of S . Then for any $N + 1$ dimensional Gaussian distribution, \hat{G} , one of whose marginals is G , there is at least one signal \hat{s} in \mathcal{S} , such that $\hat{S} = (S, \hat{s}) \sim \hat{G}$.

⁴³This forms a Markov chain: $X \rightarrow Y \rightarrow Z$.

⁴⁴For a formal definition of the chain rule see [Cover and Thomas \(2012\)](#).

Proof. Suppose $\hat{G} = \mathcal{N}(0, \hat{\Sigma})$, where $\hat{\Sigma} \in \mathbb{R}^{(N+1) \times (N+1)}$ is a positive semi-definite matrix. Since G is a marginal of \hat{G} , without loss of generality, rearrange the vectors and columns of $\hat{\Sigma}$ such that $\hat{\Sigma} = \begin{bmatrix} x & \mathbf{y}' \\ \mathbf{y} & \Sigma \end{bmatrix}$. If $x = 0$, then let $\hat{s} = 0 \sim \mathcal{N}(0, 0)$ and we are done with the proof. If not, notice that since $\hat{\Sigma}$ is positive semi-definite, its determinant has to be positive: $\det(\hat{\Sigma}) = \det(x\Sigma - \mathbf{y}\mathbf{y}') \geq 0$. Since elements of S are distinct, Σ is invertible. Also $x > 0$. We can write $\det(\hat{\Sigma}) = \det(x\Sigma) \det(\mathbf{I}_{N \times N} - x^{-1}\Sigma^{-1}\mathbf{y}\mathbf{y}') \geq 0$, which implies $\det(\mathbf{I}_{N \times N} - x^{-1}\Sigma^{-1}\mathbf{y}\mathbf{y}') = 1 - x^{-1}\mathbf{y}'\Sigma^{-1}\mathbf{y} \geq 0 \Leftrightarrow x \geq \mathbf{y}'\Sigma^{-1}\mathbf{y}$, where the equality is given by Sylvester's determinant identity. Now, choose $e_{N+1} \in \mathcal{B}$ such that $e_{N+1} \perp S$. Such an e_{N+1} exists because all the elements of S are finite linear combinations of \mathcal{B} and therefore are only correlated with a finite number of its elements, while \mathcal{B} has countably many elements.⁴⁵ Let $\hat{s} \equiv \mathbf{y}'\Sigma^{-1}\mathbf{y} + \begin{bmatrix} \sqrt{x - \mathbf{y}'\Sigma^{-1}\mathbf{y}} \\ \mathbf{0}_{N \times 1} \end{bmatrix} e_{N+1}$. Notice that $\hat{s} \in \mathcal{S}$ as it is a finite linear combination of the elements of \mathcal{B} . Notice that $\text{cov}(\hat{s}, S) = \mathbf{y}$ and $\text{var}(\hat{s}) = x$. Hence, $(\hat{s}, S) \sim \mathcal{N}(0, \hat{\Sigma})$. \square

B.3 Optimality of Linear Pricing Strategies

Every firm chooses a vector of signals $S_{j,k} \in \mathcal{S}^{n_{j,k}}$, where $n_{j,k} \in \mathbb{N}$ is the number of signals that j, k chooses to observe, and a pricing strategy $p_{j,k} : S_{j,k} \rightarrow \mathbb{R}$ that maps their signal to a price. Thus, the set of firm j, k 's pure strategies is

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

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

First, I show that in any equilibrium it has to be the case that firms' play linear pricing strategies are linear in their signals.

Lemma B.3. *Take a strategy $\zeta = (S_{j,k}, p_{j,k})_{j,k \in J \times K} \in \mathcal{A}$. Then if ζ is an equilibrium, then $\forall j, k \in J \times K$, $p_{j,k} = M'_{j,k} S_{j,k}$ for some $M_{j,k} \in \mathbb{R}^{n_{j,k}}$.*

Proof. A necessary condition for ζ to be an equilibrium is if given $(S_{j,k})_{j,k \in J \times K}$ under ζ , $\forall j, k \in J \times K$, $p_{j,k}$ solves $p_{j,k}(S_{j,k}) = \text{argmin}_{p_{j,k}} \mathbb{E}[(p_{j,k} - (1 - \alpha)q - \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l}(S_{j,l}))^2 | S_{j,k}]$. Since the objective is convex, the sufficient for minimization is if the first order condition holds: $p_{j,k}^*(S_{j,k}) = (1 - \tilde{\alpha})\mathbb{E}[q | S_{j,k}] + \tilde{\alpha}\mathbb{E}[p_j^*(S_j) | S_{j,k}]$, where $\tilde{\alpha} \equiv \frac{\alpha + \frac{\alpha}{K-1}}{1 + \frac{\alpha}{K-1}} < 1$, and $p_j^*(S_j) \equiv K^{-1} \sum_{k \in K} p_{j,k}^*(S_{j,k}^*)$. Thus, by iteration

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

⁴⁵In fact, there are countably many elements in \mathcal{B} that are orthogonal to S .

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

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

is firm j, k 's m 'th order higher order belief of its industry's average expectation of the fundamental. Similarly $\mathbb{E}_{j,k}^{(M+1)}[p_j^*(S_j)]$ is firm j, k 's $M + 1$ 'th order belief of their industry price. Assuming for now that signals are such that expectations are finite, since $\tilde{\alpha} < 1$, the later term in the limit converges to zero and we have:⁴⁶

$$p_{j,k}^*(S_{j,k}) = (1 - \tilde{\alpha}) \sum_{m=0}^{\infty} \tilde{\alpha}^m \mathbb{E}_{j,k}^{(m)}[q]. \quad (26)$$

Now, I just need to show that $\mathbb{E}_{j,k}^{(m)}[q]$ is linear in $S_{j,k}$, for all m . To see this, since all signals in \mathcal{S} are Gaussian and mean zero, $\forall j, k$, let $\Sigma_{q,S_{j,k}} \equiv \text{cov}(S_{j,k}, q) = \mathbb{E}[qS'_{j,k}]$. Also given j, k , $\forall l \neq k$, $\Sigma_{S_{j,l}, S_{j,k}} = \text{cov}(S_{j,k}, S_{j,l}) = \mathbb{E}[S_{j,l}S'_{j,k}]$ and $\Sigma_{S_{j,k}} = \text{var}(S_{j,k}) = \mathbb{E}[S_{j,k}S'_{j,k}]$.

The proof for linearity of higher order expectations is by induction: notice that for $m = 0$, $\mathbb{E}_{j,k}^{(0)}[q] = \mathbb{E}[q|S_{j,k}] = \Sigma_{q,S_{j,k}} \Sigma_{S_{j,k}}^{-1} S_{j,k}$, which implies 0'th order expectations of firms are linear in their signals. Now suppose $\forall j, l$ $\mathbb{E}_{j,l}^{(m)}[q] = A_{j,l}(m)' S_{j,l}$ for some $A_{j,l}(m) \in \mathbb{R}^{n_{j,l}}$. Now, $\mathbb{E}_{j,k}^{(m+1)}[q] = K^{-1} (A_{j,l}(m) + \underbrace{\sum_{l \neq k} A_{j,l}(m) \Sigma_{S_{j,l}, S_{j,k}} \Sigma_{S_{j,k}}^{-1}}_{\equiv A_{j,k}^{(m+1)} \in \mathbb{R}^{n_{j,k}}})' S_{j,k}$. The fact that I have assumed $\Sigma_{S_{j,k}}$ is

invertible is without loss of generality, because 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. \square

Corollary B.2. *If $\zeta = (S_{j,k} \in \mathcal{S}^{n_{j,k}}, p_{j,k}(S_{j,k}) = M'_{j,k} S_{j,k})_{j,k \in J \times K} \in \mathcal{A}$ is an equilibrium, then $\forall j, k \in J \times K$, $M_{j,k} = ((1 - \alpha) \Sigma_{q,S_{j,k}} \Sigma_{S_{j,k}}^{-1} + \alpha \frac{1}{K-1} \sum_{l \neq k} \Sigma_{S_{j,l}, S_{j,k}} \Sigma_{S_{j,k}}^{-1})'$.*

Proof. From the proof of Lemma B.3 that if ζ is an equilibrium then pricing strategies should satisfy the following optimality condition:

$$M_{j,k} S_{j,k} = (1 - \alpha) \mathbb{E}[q|S_{j,k}] + \alpha \frac{1}{K-1} \sum_{l \neq k} \mathbb{E}[M'_{j,l} S_{j,l} | S_{j,k}].$$

Thus, $M_{j,k} = ((1 - \alpha) \Sigma_{q,S_{j,k}} \Sigma_{S_{j,k}}^{-1} + \alpha \frac{1}{K-1} \sum_{l \neq k} \Sigma_{S_{j,l}, S_{j,k}} \Sigma_{S_{j,k}}^{-1})'$. \square

Given the results in this section, I restrict the set of strategies to those with linear pricing

⁴⁶if expectations are not finite, then a best response in pricing does not exist. However, since we are characterizing a necessary condition in this lemma, I characterize the best pricing responses conditional on existence.

schemes that satisfy Corollary B.2: $\mathcal{A}^* = \{\zeta \in \mathcal{A} \mid \zeta \text{ satisfies Corollary B.2}\}$.

B.4 The Attention Problem of Firms

Take a strategy $\zeta \in \mathcal{A}^*$ such that $\zeta = (S_{j,k} \in \mathcal{S}^{n_{j,k}}, p_{j,k} = M'_{j,k} S_{j,k})_{j,k \in J \times K}$. For ease of notation let $p(\zeta_{j,k}) \equiv M'_{j,k} S_{j,k}, \forall j, k \in J \times K$. Also, let $\zeta_{-(j,k)} \equiv \zeta \setminus \zeta_{j,k}$. Moreover, for any given firm $j, k \in J \times K$, let $\theta_{j,k}(\zeta_{-(j,k)}) \equiv (q, (p(\zeta_{j,l}))_{l \neq k}, (p(\zeta_{m,n}))_{m \neq j, n \in K})'$ be the augmented vector of the fundamental, the prices of other firms in j, k 's industry, and the prices of all other firms in the economy. Define $\mathbf{w} \equiv (1 - \alpha, \underbrace{\frac{\alpha}{K-1}, \dots, \frac{\alpha}{K-1}}_{K-1 \text{ times}}, \underbrace{0, 0, 0, \dots, 0}_{(J-1) \times K \text{ times}})'$. Also, for any

$\hat{\zeta}_{j,k} \in \mathcal{A}_{j,k}$, let $S(\hat{\zeta}_{j,k})$ denote the signals in \mathcal{S} that j, k observes under the strategy $\hat{\zeta}_{j,k}$. Given this notation observe that firm j, k 's problem, as defined in the text, reduces to

$$\begin{aligned} \min_{\hat{\zeta}_{j,k} \in \mathcal{A}_{j,k}} L_{j,k}(\hat{\zeta}_{j,k}, \zeta_{-(j,k)}) &\equiv \mathbb{E}[(p(\hat{\zeta}_{j,k}) - \mathbf{w}'\theta_{j,k}(\zeta_{-(j,k)}))^2 \mid S(\hat{\zeta}_{j,k})] \\ \text{s.t. } \mathcal{I}(S(\hat{\zeta}_{j,k}); \theta_{j,k}(\zeta_{-(j,k)})) &\leq \kappa, \end{aligned} \quad (27)$$

where given the joint distribution of $(S(\hat{\zeta}_{j,k}), \theta_{j,k}(\zeta_{-(j,k)}))$, the mutual information is defined in Section B.1. It is also useful to restate the definition of the equilibrium given this notation:

Definition. An equilibrium is a strategy $\zeta \in \mathcal{A}$ such that $\forall j, k \in J \times K$

$$\zeta_{j,k} = \operatorname{argmin}_{\zeta'_{j,k} \in \mathcal{A}_{j,k}} L_{j,k}(\zeta'_{j,k}, \zeta_{-(j,k)}) \text{ s.t. } \mathcal{I}(S(\zeta_{j,k}); \theta_{j,k}(\zeta_{-(j,k)})) \leq \kappa. \quad (28)$$

The solution to this problem, if exists, is not unique. To show this, I define the following relation on the deviations of j, k , given a strategy $\zeta \in \mathcal{A}^*$, and show that it is an equivalence.

Definition. For any two distinct elements $\{\zeta^1_{j,k}, \zeta^2_{j,k}\} \subset \mathcal{A}_{j,k}$, and given $\zeta = (\zeta_{j,k}, \zeta_{-(j,k)}) \in \mathcal{A}^*$, we say $\zeta^1_{j,k} \sim_{j,k \mid \zeta} \zeta^2_{j,k}$ if $L_{j,k}(\zeta^1_{j,k}, \zeta_{-(j,k)}) = L_{j,k}(\zeta^2_{j,k}, \zeta_{-(j,k)})$, where $L_{j,k}(\cdot, \cdot)$ is defined as in Equation (27). Note that $\forall j, k \in J \times K$ and $\forall \zeta \in \mathcal{A}^*$, $\sim_{j,k \mid \zeta}$ is an equivalence relation as reflexivity, symmetry and transitivity are trivially satisfied by properties of equality.

By definition the agent is indifferent between elements of an equivalence class. Now, given $\zeta = (\zeta_{j,k}, \zeta_{-(j,k)}) \in \mathcal{A}^*$, let $[\hat{\zeta}_{j,k}]_{\zeta} \equiv \{\zeta'_{j,k} \in \mathcal{A}_{j,k} \mid \zeta'_{j,k} \sim_{j,k \mid \zeta} \hat{\zeta}_{j,k}\}$. The following lemma shows there is always a deviation with a single dimensional signal that requires less attention but yields the same payoff. Therefore, for any strategy of others, optimal signal choice of a firm is one dimensional.

Lemma B.4. For any $j, k \in J \times K, \forall \zeta = (\zeta_{j,k}, \zeta_{-(j,k)}) \in \mathcal{A}^*, \exists \hat{\zeta}_{j,k} \in [\zeta_{j,k}]_{\zeta}$ such that the agent observes only one signal under $\hat{\zeta}_{j,k}$ and $\mathcal{I}(S(\hat{\zeta}_{j,k}); \theta_{j,k}(\zeta_{-(j,k)})) \leq \mathcal{I}(S(\zeta_{j,k}); \theta_{j,k}(\zeta_{-(j,k)}))$.

Moreover, $\hat{\varsigma}_{j,k}$ does not alter the covariance of firm j, k 's price with the fundamental and the prices of all other firms in the economy under ς .

Proof. I prove this lemma by constructing such an strategy. Given $\varsigma \in \mathcal{A}^*$, let $\Sigma_{\varsigma_{j,k}} \equiv \text{var}(S(\varsigma_{j,k}))$, $\Sigma_{\theta_{j,k}, \varsigma_{j,k}} \equiv \text{cov}(\theta_{j,k}(\varsigma_{-(j,k)}), S(\varsigma_{j,k}))$ and $\Sigma_{\theta_{j,k}} \equiv \text{var}(\theta_{j,k}(\varsigma_{-(j,k)}))$. Thus,

$$(S(\varsigma_{j,k}), \theta_{j,k}(\varsigma_{-(j,k)})) \sim \mathcal{N}(0, \begin{bmatrix} \Sigma_{\varsigma_{j,k}} & \Sigma'_{\theta_{j,k}, \varsigma_{j,k}} \\ \Sigma_{\theta_{j,k}, \varsigma_{j,k}} & \Sigma_{\theta_{j,k}} \end{bmatrix}).$$

Moreover, since $\varsigma \in \mathcal{A}^*$, then pricing strategies are linear, and by Corollary B.2 $p_{j,k}(\varsigma) = \mathbf{w}' \mathbb{E}[\theta_{j,k}(\varsigma_{-(j,k)}) | S(\varsigma_{j,k})] = \mathbf{w}' \Sigma_{\theta_{j,k}, \varsigma_{j,k}} \Sigma_{\varsigma_{j,k}}^{-1} S(\varsigma_{j,k})$. Notice that

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

Now, let $\hat{\varsigma}_{j,k} \equiv \mathbf{w}' \Sigma_{\theta_{j,k}, \varsigma_{j,k}} \Sigma_{\varsigma_{j,k}}^{-1} S(\varsigma_{j,k})$. Clearly, $\hat{\varsigma}_{j,k} \in \mathcal{S}$ as it is a finite linear combination of the elements of $S_{j,k}$, and \mathcal{S} is rich. Define $\hat{\varsigma}_{j,k} \equiv (\hat{\varsigma}_{j,k}, 1) \in \mathcal{A}_{j,k}$. Notice that

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

Thus, $\hat{\varsigma}_{j,k} \in [S_{j,k}]_{\varsigma}$. Also, observe that $\theta_{j,k}(\varsigma_{-(j,k)}) \perp \hat{\varsigma}_{j,k} | S(\varsigma_{j,k})$. Therefore, by the data processing inequality in Lemma B.1, $\mathcal{I}(\hat{\varsigma}_{j,k}; \theta_{j,k}(\varsigma_{-(j,k)})) \leq \mathcal{I}(S(\varsigma_{j,k}); \theta_{j,k}(\varsigma_{-(j,k)}))$. Finally, observe that $p_{j,k}(\hat{\varsigma}_{j,k}, \varsigma_{-(j,k)}) = p_{j,k}(\varsigma) = \mathbf{w}' \Sigma_{\theta_{j,k}, \varsigma_{j,k}} \Sigma_{\varsigma_{j,k}}^{-1} S(\varsigma_{j,k})$. Thus, the covariance of j, k 's price with all the elements of $\theta_{j,k}(\varsigma_{-(j,k)})$ remains unchanged when j, k deviates from $\varsigma_{j,k}$ to $\hat{\varsigma}_{j,k}$. \square

B.5 Equilibrium Signals

Let $\mathcal{E} \equiv \{\varsigma \in \mathcal{A} | \varsigma \text{ is an equilibrium as stated in Statement (28)}\}$ denote the set of equilibria for the game. The following definition states an equivalence relation among the equilibria.

Definition. Suppose $\{\varsigma_1, \varsigma_2\} \subset \mathcal{E}$. We say $\varsigma_1 \sim_{\mathcal{E}} \varsigma_2$ if they imply the same joint distribution for prices of firms and the fundamental. Formally, $\varsigma_1 \sim_{\mathcal{E}} \varsigma_2$ if given that $(q, p_{j,k}(\varsigma_1))_{j,k \in J \times K} \sim G$, then $(q, p_{j,k}(\varsigma_2))_{j,k \in J \times K} \sim G$ as well. This is trivially an equivalence relation as it satisfies reflexivity, symmetry and transitivity by properties of equality.

Lemma B.5. Let $\mathcal{A}^{**} \equiv \{\varsigma \in \mathcal{A} | \varsigma = (s_{j,k} \in \mathcal{S}, 1)_{j,k \in J \times K}\}$. Suppose $\varsigma \in \mathcal{A}$ is an equilibrium for the game. Then, there exists $\hat{\varsigma} \in \mathcal{A}^{**}$ such that $\hat{\varsigma} \sim_{\mathcal{E}} \varsigma$.

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 loop for all firms: we

know firm 1,1 has a strategy $\hat{\varsigma}_{1,1} = (s_{1,1} \in \mathcal{S}, 1)$ 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 by the previous lemma $\hat{\varsigma}_{1,1}$ does not alter the joint distribution of prices. 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 is the same under both strategies. Now, repeat the same thing 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 $\varsigma^{J,K} = (\hat{\varsigma}_{j,k})_{j,k \in J \times K}$, which is (1) an equilibrium and (2) implies the same joint distribution among prices and fundamentals as ς . Moreover, notice that $\varsigma^{J,K} \in \mathcal{A}^{**}$. \square

So far we have shown that any equilibrium has an equivalent in \mathcal{A}^{**} , so as long as we are interested in the joint distribution of prices and the fundamental it suffices to only look at equilibria in this set. The next lemma shows that given any strategy $\varsigma \in \mathcal{A}^{**}$, for any $j, k \in J \times K$, the set of j, k 's deviations is equivalent to choosing a joint distribution between their price and $\theta_{j,k}(\varsigma_{-(j,k)})$.

Lemma B.6. *Suppose $\varsigma \in \mathcal{A}^{**}$ is an equilibrium. Then, $\forall j, k \in J \times K$, any deviation for j, k is equivalent to a Gaussian joint distribution between their price and $\theta_{j,k}(\varsigma_{-(j,k)})$. Moreover, if two different deviations of j, k imply the same joint distribution for prices and the fundamental, they both require the same amount of attention and the firm is indifferent between.*

Proof. Given ς , let $\Sigma_{\theta_{j,k}}$ be such that $\theta_{j,k}(\varsigma_{-(j,k)}) \sim \mathcal{N}(0, \Sigma_{\theta_{j,k}})$. Notice that $\Sigma_{\theta_{j,k}}$ has to be invertible: if not, then there must a firm whose signal is either co-linear with the fundamental or the signal of another firm, meaning that their signal is perfectly correlated with one of those. But that violates the capacity constraint of that firm as they are processing infinite capacity, which is a contradiction with the assumption that ς is an equilibrium.⁴⁷

Now, from Lemma B.4 we know that it suffices to look at deviations of the form $(s_{j,k} \in \mathcal{S}, 1)$. First, observe that any deviation of the firm j, k creates a Gaussian joint distribution for $(s_{j,k}, \theta_{j,k}(\varsigma_{-(j,k)}))$ as $s_{j,k} \in \mathcal{S}$. Moreover, suppose $G = \mathcal{N}(0, \begin{bmatrix} x & \mathbf{y}' \\ \mathbf{y} & \Sigma_{\theta_{j,k}} \end{bmatrix})$ is a Gaussian distribution. Since $\Sigma_{\theta_{j,k}}$ is invertible, Corollary B.1 implies that there is a signal $s_{j,k} \in \mathcal{S}$, such that $(s_{j,k}, \theta_{j,k}(\varsigma_{-(j,k)})) \sim G$.

For the last part of the lemma, suppose for two different signals $s_{j,k}^1$ and $s_{j,k}^2$ in \mathcal{S} , $(s_{j,k}^1, \theta_{j,k}(\varsigma_{-(j,k)}))$

⁴⁷Recall, for any two one dimensional Normal random variables X and Y , $I(X, Y) = -\frac{1}{2} \log_2(1 - \rho_{X,Y}^2)$, where $\rho_{X,Y}$ is the correlation of X and Y . Notice that $\lim_{\rho^2 \rightarrow 1} I(X, Y) \rightarrow +\infty$.

and $(s_{j,k}^2, \theta_{j,k}(\zeta_{-(j,k)}))$ have the same joint distribution. Then,

$$\text{var}(\theta_{j,k}(\zeta_{-(j,k)})|s_{j,k}^1) = \text{var}(\theta_{j,k}(\zeta_{-(j,k)})|s_{j,k}^2)$$

which implies that $L_{j,k}((s_{j,k}^1, 1), \zeta_{-(j,k)}) = L_{j,k}((s_{j,k}^2, 1), \zeta_{-(j,k)})$. Moreover, given that the conditional variances under both signals are the same we have

$$\mathcal{I}(s_{j,k}^1; \theta_{j,k}(\zeta_{-(j,k)})) = \mathcal{I}(s_{j,k}^2; \theta_{j,k}(\zeta_{-(j,k)})).$$

Therefore, the firm is indifferent between $s_{j,k}^1$ and $s_{j,k}^2$. \square

This last lemma ensures us that instead of considering all the possible deviations in \mathcal{S} , we can look among all the possible joint distributions. If there is a joint distribution that solves a firm's problem, then the lemma implies that there is a signal in the set of available signals that creates that joint distribution.

Lemma B.7. *Suppose $\zeta = (s_{j,k}^* \in \mathcal{S}, 1) \in \mathcal{A}^{**}$ is an equilibrium, then $\forall j, k \in J \times K$,*

$$s_{j,k}^* = \lambda \mathbf{w}' \theta_{j,k}(\zeta_{-(j,k)}) + z_{j,k}, \quad z_{j,k} \perp \theta_{j,k}(\zeta_{-(j,k)}), \quad \text{var}(z_{j,k}) = \lambda(1 - \lambda) \text{var}(\mathbf{w}' \theta_{j,k}(\zeta_{-(j,k)})).$$

Proof. For firm $j, k \in J \times K$, let $\Sigma_{\theta_{j,k}}$ denote the covariance matrix of $\theta_{j,k}(\zeta_{-(j,k)})$. From Lemma B.4 it is sufficient to look at deviations of the form $(s_{j,k} \in \mathcal{S}, 1)$. For a given $s_{j,k} \in \mathcal{S}$, $(s_{j,k}, \theta_{j,k}(\zeta_{-(j,k)})) \sim \mathcal{N}(0, \Sigma_{s_{j,k}, \theta_{j,k}})$, where $\Sigma_{s_{j,k}, \theta_{j,k}} = \begin{bmatrix} x^2 & \mathbf{y}' \\ \mathbf{y} & \Sigma_{\theta_{j,k}} \end{bmatrix} \succeq 0$. First, recall that for $(s_{j,k} \in \mathcal{S}, 1)$ to be optimal, it has to be the case that $p_{j,k} = \mathbf{w}' \mathbb{E}[\theta_{j,k}(\zeta_{-(j,k)})|s_{j,k}] = x^{-2} \mathbf{w}' \mathbf{y}_{s_{j,k}}$. Thus,

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

Now, given $s_{j,k} \in \mathcal{S}$, the firm's loss in profits is $\text{var}(\mathbf{w}' \theta_{j,k}(\zeta_{-(j,k)})|s_{j,k}) = \mathbf{w}' \Sigma_{\theta_{j,k}} \mathbf{w} - x^{-2} (\mathbf{w}' \mathbf{y})^2$ and the capacity constraint is $\frac{1}{2} \log_2(|\mathbf{I} - x^{-2} \Sigma_{\theta_{j,k}}^{-1} \mathbf{y} \mathbf{y}'|) \geq -\kappa \Leftrightarrow x^{-2} \mathbf{y}' \Sigma_{\theta_{j,k}}^{-1} \mathbf{y} \leq \lambda \equiv 1 - 2^{-2\kappa}$.

Moreover, from the previous lemma we know that for any (x, \mathbf{y}) such that $\begin{bmatrix} x^2 & \mathbf{y}' \\ \mathbf{y} & \Sigma_{\theta_{j,k}} \end{bmatrix} \succeq 0$, then there is a signal in \mathcal{S} that creates this joint distribution. Therefore, we let the agent choose (x, \mathbf{y}) freely to solve $\min_{(x, \mathbf{y})} \mathbf{w}' \Sigma_{\theta_{j,k}} \mathbf{w} - x^{-2} (\mathbf{w}' \mathbf{y})^2$ s.t. $x^{-2} \mathbf{y}' \Sigma_{\theta_{j,k}}^{-1} \mathbf{y} \leq \lambda$. The solution can be derived by taking first order conditions, but there is simpler a way. Notice that by Cauchy-Schwarz inequality $x^{-2} (\mathbf{w}' \mathbf{y})^2 = x^{-2} (\Sigma_{\theta_{j,k}}^{\frac{1}{2}} \mathbf{w})' (\Sigma_{\theta_{j,k}}^{-\frac{1}{2}} \mathbf{y}) \leq x^{-2} (\mathbf{w}' \Sigma_{\theta_{j,k}} \mathbf{w}) (\mathbf{y}' \Sigma_{\theta_{j,k}}^{-1} \mathbf{y})$. Therefore,

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

where, the last line is from the capacity constraint. This defines a global lower-bound for the objective of the firm that holds for any choice of (x, \mathbf{y}) . However, this global minimum is attained if both the Cauchy-Schwarz inequality and the capacity constraint bind. From the properties of the Cauchy-Schwarz inequality, we know it binds 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, the capacity constraint binds if $c_0 = \sqrt{\frac{\lambda}{\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w}}}$. 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, to find a signal that creates this joint distribution, choose $s_{j,k}^* \in \mathcal{S}$ such that

$$s_{j,k}^* = \lambda\mathbf{w}'\theta_{j,k}(\zeta_{-(j,k)}) + z_{j,k}, z_{j,k} \perp \theta_{j,k}(\zeta_{-(j,k)}), \text{var}(z_{j,k}) = \lambda(1 - \lambda)\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w}.$$

notice that $\text{cov}(s_{j,k}^*, \theta_{j,k}(\zeta_{-(j,k)})) = \lambda\Sigma_{\theta_{j,k}}\mathbf{w}$, and $\text{var}(s_{j,k}^*) = \lambda\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w}$. Notice that this implies the equilibrium set of signals are

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

where $\text{var}(z_{j,t}) = \lambda(1 - \lambda)\text{var}((1 - \alpha)q + \alpha\frac{1}{K-1}\sum_{l \neq k} s_{j,l}^*)$. □

B.6 Uniqueness of Equilibrium in Joint Distribution of Prices

Having specified the equilibrium signals, I now show that all equilibria imply the same joint distribution of prices.

Lemma B.8. *Suppose $\alpha \in [0, 1)$. Then, $\mathcal{E} / \sim_{\mathcal{E}}$ is non-empty and a singleton.*

Proof. I show this by directly characterizing the equilibrium. From previous section we know that any equilibrium is equivalent to one in strategies of \mathcal{A}^{**} . Suppose that $(s_{j,k}^*, 1)_{j,k \in J \times K} \in \mathcal{A}^{**}$ is an equilibrium, and notice that in this equilibrium every firm simply sets their price equal to their signal, $p_{j,k} \equiv s_{j,k}^*$. Also, Lemma 8 showed that this equilibrium signals should satisfy the following

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

where $\text{var}(z_{j,t}) = \lambda(1 - \lambda)\text{var}((1 - \alpha)q + \alpha\frac{1}{K-1}\sum_{l \neq k} p_{j,l})$. 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.

I 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 straight forward to show that $cov(p_j, q) = \frac{\lambda - \lambda\alpha}{1 - \lambda\alpha}\mathbf{1}$. Thus, in any equilibrium, the covariance of any firm's price with the fundamental q has to be equal to $\delta \equiv \frac{\lambda - \lambda\alpha}{1 - \lambda\alpha}$.

Next, I show that for any two firms in two different industries, their prices 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 $z_j \perp p_{l,m}$. Now, notice that

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

With some algebra, we get $cov(p_j, p_{l,m}) = \delta^2\mathbf{1} \Rightarrow cov(p_j, p_{l,m}|q) = 0$. 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 specify 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 $\Sigma_{z,j} = cov(z_j, z_j)$ and $\Sigma_{p,j} = cov(p_j, p_j)$. 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 $\mathbf{D}_j \equiv cov(p_j, z_j) = \mathbf{B}\Sigma_{z,j}$ where \mathbf{D}_j is a diagonal matrix whose k 'th element on the diagonal is $var(z_{j,k})$. From the equilibrium conditions we have

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

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 tedious algebra, we get that the implied distribution for prices is such that

$$var(p_{j,k}) = \frac{1 - \alpha\lambda}{1 - \alpha\tilde{\lambda}}\lambda^{-1}\delta^2, \forall j, k; \quad cov(p_{j,k}, p_{j,l}) = \frac{1 - \alpha\lambda}{1 - \alpha\tilde{\lambda}}\frac{\tilde{\lambda}}{\lambda}\delta^2, \forall j, k, l \neq k,$$

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

B.7 Reinterpretation of a Firm's Attention Problem.

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: $cov(p_{j,k}, p_{l,m}|q)_{l \neq j} = 0$. Now, take strategy $\varsigma_{j,-k}$ for other firms and decompose the average price of others such that $p_{j,-k}(\varsigma_{j,-k}) = \frac{1}{K-1} \sum_{l \neq k} p_{j,l}(\varsigma_{j,l}) = \delta q + v_{j,-k}$, where δ and the joint $var(v_{j,-k})$ is implied by $\varsigma_{j,-k}$. Let $\sigma_v^2 \equiv var(v_{j,-k})$ be the variance of the average mistake of other firms in j, k 's industry when they play the strategy. For $s_{j,k} \in \mathcal{S}$, and 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 they observe $s_{j,k}$ is

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

With some algebra, it is straight forward to show that

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

Now, to derive the information constraint in terms of the two correlations: $\mathcal{I}(s_{j,k}; (q, p_{j,-k}^*)) \leq \kappa \Leftrightarrow \frac{1}{2} \log_2(\frac{var(s_j)}{var(s_{j,k}|(q, p_{j,-k}^*))}) \leq \kappa$. Notice that $\frac{var(s_j|q, p_{j,-k}^*)}{var(s_j)} = 1 - (\rho_q(s_j)^2 + \rho_v(s_j)^2)$. Thus, the information constraint becomes $\rho_q^2(s_j) + \rho_v^2(s_j) \leq \lambda \equiv 1 - 2^{-2\kappa}$. So j, k 's problem reduces to

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

The problem reduces to choosing correlations as the information set is rich: for any pair of $(\rho_q, \rho_v) \in [-1, 1]^2$, there is a signal in \mathcal{S} that generates that pair.

B.8 Proofs of Propositions for the Static Model

Here I include the proofs of Propositions 1 to 3. The proofs and derivations for Section 4 are included in Appendix F.

Proof of Proposition 1.

1. Given the result in Lemma B.8, notice that since attention is strictly increasing in the squared correlation: $\rho_q^{*2} = \frac{cov(p_{j,k}, q)^2}{var(p_{j,k})} = \frac{K-1+\alpha\delta}{K-1+\alpha\lambda} \lambda$. However, 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, $\rho_v^{*2} = \lambda - \rho_q^{*2} > 0$, meaning that firms pay attention to the mistakes of their competitors.

2. From the previous part, notice that $\frac{\partial \rho_q^{*2}}{\partial K} \frac{1}{\rho_q^{*2}} = \frac{\alpha(\lambda - \delta)}{(K-1+\alpha\lambda)(K-1+\alpha\delta)} > 0$. Also

$$\frac{\partial \rho_q^{*2}}{\partial \alpha} \frac{1}{\rho_q^{*2}} = \frac{(K-1)(\delta - \lambda) + (K-1 + \alpha\lambda)\alpha \frac{\partial \delta}{\partial \alpha}}{(K-1 + \alpha\delta)(K-1 + \alpha\lambda)} < 0.$$

The inequality comes from $\delta - \lambda < 0$ and $\frac{\partial \delta}{\partial \alpha} = \delta \frac{\lambda - 1}{(1-\alpha)(1-\alpha\lambda)} < 0$.

3. Shown in the proof of Lemma B.8.

Proof of Proposition 2.

First of all notice 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, the average converges to zero by the law of large numbers as $J \rightarrow \infty$. Therefore, $p = \delta q$. Moreover, $\mathbb{E}^{j,k}[p_{j,-k}] = \frac{\text{cov}(s_{j,k}, p_{j,-k})}{\text{var}(p_{j,k})} s_{j,k} = \tilde{\lambda} p_{j,k}$ and $\mathbb{E}^{j,k}[p] = \frac{\text{cov}(s_{j,k}, p)}{\text{var}(p_{j,k})} p_{j,k} = \frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda} \lambda p_{j,k}$ where $\tilde{\lambda} = \frac{\lambda(K-1)+\alpha\lambda}{K-1+\alpha\lambda} > \lambda$ is defined as in the proof of Lemma B.8. So, $\overline{\mathbb{E}^{j,k}[p_{j,-k}]} = \tilde{\lambda} p$, $\overline{\mathbb{E}^{j,k}[p]} = \frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda} \lambda p$. Finally,

$$\text{cov}(\overline{\mathbb{E}^{j,k}[p_{j,-k}]}, p) = \tilde{\lambda} \text{var}(p) > \frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda} \lambda \text{var}(p) = \text{cov}(\overline{\mathbb{E}^{j,k}[p]}, p).$$

Also, if $K \rightarrow \infty$ then $\tilde{\lambda} \rightarrow \lambda$ and $\text{cov}(\overline{\mathbb{E}^{j,k}[p]}, p) \rightarrow \text{cov}(\overline{\mathbb{E}^{j,k}[p_{j,-k}]}, p)$.

Proof of Corollary 1.

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

C Available Information in the Dynamic Model

The set of available signals in the dynamic model is an extension of the one defined in Appendix B.2. The main difference is the notion of time and the fact that at every period nature draws new shocks and the set of the available information in the economy expands. To capture this evolution, I define a signal structure as a sequence of sets $(\mathcal{S}^t)_{t=-\infty}^{\infty}$ where $\mathcal{S}^{t-s} \subset \mathcal{S}^t, \forall s \geq 0$. Here, \mathcal{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 at every period, in addition to the shock to the nominal demand, the nature draws countably infinite uncorrelated standard normal noises. Similar to Appendix B.2, let \mathcal{S}_t be the set of all finite linear combinations of these uncorrelated noises. Now, define $\mathcal{S}^t = \{\sum_{s=0}^{\infty} a_s e_{t-s} | \forall \tau \geq 0, a_\tau \in \mathbb{R}, e_{t-\tau} \in \mathcal{S}_{t-\tau}\}, \forall t$. First of all, notice that for all $t, q_t \in \mathcal{S}^t$, as it is a linear combination of all $u_{t-\tau}$'s and $u_{t-\tau} \in \mathcal{S}_{t-\tau}, \forall \tau \geq 0$. This implies that perfect information is available about the fundamentals of the economy.

D Derivations

Solution to Household's Problem (7).

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 $\mathcal{C}_{j,t} \equiv (C_{j,1,t}, \dots, C_{j,K,t})$ be the vector of household's consumption from firms in industry $j \in J$, so that $C_{j,t} \equiv \Phi(\mathcal{C}_{j,t})$ where $\Phi(\cdot)$ 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 as well as the Kimball aggregator discussed in Appendix E. Moreover, for less crowded notation, I refer to K_j as K whenever the industry index is redundant. First, I derive the demand of the household for different goods. $\forall j, k \in J \times K$ the first order condition with respect to $C_{j,k,t}$ is

$$P_{j,k,t} = \frac{1}{J} \frac{\varphi_{2,t}}{\varphi_{1,t}} C_t \frac{\Phi_k(\mathcal{C}_{j,t})}{\Phi(\mathcal{C}_{j,t})} \quad (29)$$

where $\Phi_k(\mathcal{C}_{j,t}) \equiv \frac{\partial \Phi(\mathcal{C}_{j,t})}{\partial C_{j,k,t}}$. Notice that given these optimality conditions $\sum_{(j,k) \in J \times K} P_{j,k,t} C_{j,k,t} = \frac{1}{J} \frac{\varphi_{2,t}}{\varphi_{1,t}} C_t \sum_{j \in J} \underbrace{\sum_{k \in K} \frac{\Phi_k(\mathcal{C}_{j,t})}{\Phi(\mathcal{C}_{j,t})} C_{j,k,t}}_{=1, \forall j \in J} = \frac{\varphi_{2,t}}{\varphi_{1,t}} C_t$, where the equality under curly bracket is from Euler

theorem for homogeneous functions as $\Phi(\cdot)$ is CRS. Therefore, $P_t \equiv \frac{\varphi_{2,t}}{\varphi_{1,t}}$ is the price of the aggregate consumption basket C_t . Now, from Equation (29), $\mathcal{P}_{j,t} \equiv (P_{j,1,t}, \dots, P_{j,K,t}) = \nabla \log(\Phi(\frac{C_{j,t}}{J^{-1} P_t C_t}))$. I need to show that this function is invertible to prove that a demand function exists. For ease of notation, define function $f : \mathbb{R}^K \rightarrow \mathbb{R}^K$ such that $f(\mathbf{x}) \equiv \nabla \log(\Phi(\mathbf{x}))$. Notice that $f(\cdot)$ 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 $\mathcal{J}_{m,n}^f(\mathbf{x}) \equiv \frac{\partial}{\partial x_n} \frac{\Phi_m(\mathbf{x})}{\Phi(\mathbf{x})} = \frac{\Phi_{m,n}(\mathbf{x})}{\Phi(\mathbf{x})} - \frac{\Phi_n(\mathbf{x}) \Phi_m(\mathbf{x})}{\Phi(\mathbf{x})^2}$. Let $\mathbf{1}$ be the unit vector in \mathbb{R}^K . Since $\Phi(\cdot)$ is symmetric along its arguments, for any $k \in (1, \dots, K)$, $\Phi_1(\mathbf{1}) = \Phi_k(\mathbf{1})$, $\Phi_{11}(\mathbf{1}) = \Phi_{kk}(\mathbf{1}) < 0$. Since $\Phi(\cdot)$ 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 $\Phi_k(\cdot)$ 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 $l \neq k$, $\Phi_{kl}(\mathbf{1}) = -\frac{1}{K-1}\Phi_{1l}(\mathbf{1}) > 0$. This last equation implies that $\mathcal{J}^f(\mathbf{1})$ is an invertible matrix.⁴⁹ Therefore, by inverse function theorem $f(\cdot)$ 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 $\frac{C_{j,t}}{J^{-1}P_t C_t} = f^{-1}(\mathcal{P}_{j,t})$. It is straight forward to show that $f^{-1}(\cdot)$ 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, $C_{j,k,t} = J^{-1}P_t C_t f_k^{-1}(\mathcal{P}_{j,t})$, where $f_k^{-1}(\mathbf{x})$ is the k 'th element of the vector $f^{-1}(\mathcal{P}_{j,t})$. Finally, since $f(\cdot)$ 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: $C_{j,k,t} = P_t C_t \mathcal{D}(P_{j,k,t}, P_{j,-k,t})$, where $\mathcal{D}(\cdot, \cdot)$ 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 $P_t C_t = \beta(1 + i_t)\mathbb{E}_t^f [P_{t+1} C_{t+1}]$ and $P_t C_t = W_t$.

Loss Function of the Firms.

Let $\Pi(P_{j,k,t}, P_{j,-k,t}, W_t) = (P_{j,k,t} - (1 - \bar{s})W_t)\mathcal{D}(P_{j,k,t}, P_{j,-k,t})$ denote the profit function of the firm following the text. Notice that this function is homogeneous of degree 1 as $\mathcal{D}(\cdot, \cdot)$ is homogeneous of degree -1. Now for any given set of signals over time that firm j, k could choose to see, its profit maximization problem is

$$\max_{(P_{j,k,t}: S_{j,k}^t \rightarrow \mathbb{R})_{t=0}^{\infty}} \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t Q_0 \Pi(P_{j,k,t}, P_{j,-k,t}, W_t) | S_{j,k}^{-1} \right].$$

Define the loss function of firm from mispricing at a certain time as

$$L(P_{j,k,t}, P_{j,-k,t}, W_t) \equiv \Pi(P_{j,k,t}^*, P_{j,-k,t}, W_t) - \Pi(P_{j,k,t}, P_{j,-k,t}, W_t),$$

where $P_{j,k,t}^* = \operatorname{argmax}_x \Pi(x, P_{j,-k,t}, W_t)$. Note that

$$\min_{(P_{j,k,t}: S_{j,k}^t \rightarrow \mathbb{R})_{t=0}^{\infty}} \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t Q_0 L(P_{j,k,t}, P_{j,-k,t}, W_t) | S_{j,k}^{-1} \right]$$

⁴⁸Follows 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 $\Phi_k(a\mathbf{x}) = \Phi_k(\mathbf{x})$.

⁴⁹With 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.

has the same solution as profit maximization problem of the firm because $L(\cdot)$ is also homogeneous of degree 1 and $\sum_{t=0}^{\infty} \beta^t \frac{Q_0}{Q_t} \max_x \Pi(x, P_{j,-k,t}, W_t)$ is independent of $(P_{j,k,t})_{t=0}^{\infty}$. Now, I take a second order approximation to

$$\mathcal{L}[(P_{j,k,t}, P_{j,-k,t}, Q_t, W_t)_{t=0}^{\infty}] \equiv \sum_{t=0}^{\infty} \beta^t Q_0 L(P_{j,k,t}, P_{j,-k,t}, W_t)$$

around a symmetric point where $\forall t, P_{j,k,t} = P_{j,l,t} |_{\forall l \neq k} = \bar{P}, W_t = \bar{Q}$ such that $\bar{P} = \operatorname{argmax}_x \Pi(x, \bar{P}, \bar{Q})$. For any of variables above let its corresponding small letter denote percentage deviation of that variable from this symmetric point ($q_t \equiv \frac{Q_t - \bar{Q}}{\bar{Q}}$ and so on). Observe that up to second order terms

$$\begin{aligned} L(P_{j,k,t}, P_{j,-k,t}, W_t) &\approx L(\bar{P}, \bar{P}, \bar{Q}) + (p_{j,k,t}^* - p_{j,k,t}) \bar{P} \frac{\partial}{\partial P_{j,k,t}} \Pi(\bar{P}, \bar{P}, \bar{Q}) \\ &+ (p_{j,k,t}^{*2} - p_{j,k,t}^2) \frac{\bar{P}^2}{2} \frac{\partial^2}{\partial P_{j,k,t}^2} \Pi(\bar{P}, \bar{P}, \bar{Q}) \\ &+ (p_{j,k,t}^* - p_{j,k,t}) \sum_{l \neq k} p_{j,l,t} \bar{P}^2 \frac{\partial^2}{\partial P_{j,k,t} \partial P_{j,l,t}} \Pi(\bar{P}, \bar{P}, \bar{Q}) \\ &+ (p_{j,k,t}^* - p_{j,k,t}) w_t \bar{Q} \bar{P} \frac{\partial^2}{\partial P_{j,k,t} \partial W_t} \Pi(\bar{P}, \bar{P}, \bar{Q}). \end{aligned}$$

But notice that $L(\bar{P}, \bar{P}, \bar{Q}) = 0$, and $p_{j,k,t}^* = \frac{P_{j,k,t}^* - \bar{P}}{\bar{P}}$ is such that $\Pi_1(P_{j,k,t}^*, P_{j,-k,t}, Q_t) = 0$, meaning that

$$p_{j,k,t}^* \bar{P} \frac{\partial^2 \Pi(\bar{P}, \bar{P}, \bar{Q})}{\partial P_{j,k,t}^2} + \sum_{l \neq k} p_{j,l,t} \bar{P} \frac{\partial^2 \Pi(\bar{P}, \bar{P}, \bar{Q})}{\partial P_{j,k,t} \partial P_{j,l,t}} + w_t \bar{Q} \frac{\partial^2 \Pi(\bar{P}, \bar{P}, \bar{Q})}{\partial P_{j,k,t} \partial W_t} = 0.$$

Plug this into the above approximation to get $L(P_{j,k,t}, P_{j,-k,t}, W_t) = -\frac{\bar{P}^2}{2} \Pi_{11}(p_{j,k,t} - p_{j,k,t}^*)^2$. Therefore, the approximation gives

$$\mathcal{L}[(P_{j,k,t}, P_{j,-k,t}, Q_t, W_t)_{t=0}^{\infty}] \approx \underbrace{-\frac{1}{2} \Pi_{11} \bar{Q} \bar{P}^2}_{>0} \sum_{t=0}^{\infty} \beta^t (p_{j,k,t} - p_{j,k,t}^*)^2,$$

which implies that up to this second order approximation the profit maximization of the firm is equivalent to $\min_{(p_{j,k,t}: S_{j,k}^t \rightarrow \mathbb{R})_{t=0}^{\infty}} \mathbb{E}[\sum_{t=0}^{\infty} \beta^t (p_{j,k,t} - p_{j,k,t}^*)^2 | S_{j,k}^{-1}]$.

General Form of α .

To derive the expression for $p_{j,k,t}^*$, recall that $P_{j,k,t}^*$ is such that $\Pi_1(P_{j,k,t}^*, P_{j,-k,t}, W_t) = 0$. Considering the specific form of the profit function this gives $P_{j,k,t}^* = \frac{\varepsilon_D(P_{j,k,t}^*, P_{j,-k,t})}{\varepsilon_D(P_{j,k,t}^*, P_{j,-k,t}) - 1} (1 - \bar{s}) Q_t$ where $\varepsilon_D(P_{j,k,t}^*, P_{j,-k,t}) \equiv -\frac{\partial \mathcal{D}(P_{j,k,t}^*, P_{j,-k,t})}{\partial P_{j,k,t}^*} \frac{P_{j,k,t}^*}{\mathcal{D}(P_{j,k,t}^*, P_{j,-k,t})}$. Define the super-elasticity of demand for a firm as $\varepsilon_D^\varepsilon(P_{j,k,t}, P_{j,-k,t}) \equiv \frac{P_{j,k,t}}{\varepsilon_D(P_{j,k,t}, P_{j,-k,t})} \frac{\partial}{\partial P_{j,k,t}} \varepsilon_D(P_{j,k,t}, P_{j,-k,t})$. Since $\mathcal{D}(\cdot, \cdot)$ is homogeneous of degree -1, then $\varepsilon_D(\cdot, \cdot)$ and $\varepsilon_D^\varepsilon(\cdot, \cdot)$ are both homogeneous of degree zero. For ease of notation let $\varepsilon_D \equiv \varepsilon_D(1, 1)$ and $\varepsilon_D^\varepsilon \equiv \varepsilon_D^\varepsilon(1, 1)$. Now, recall from the previous section that $p_{j,k,t}^*$ is derived by a first order log-linearization of this equation, which implies $p_{j,k,t}^* = (1 - \alpha)q_t + \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l,t}$, where

$$\alpha \equiv \frac{\varepsilon_D^\varepsilon}{\varepsilon_D^\varepsilon + \varepsilon_D - 1}. \quad (30)$$

Notice that $\alpha \in [0, 1)$ as long as $\varepsilon_D^\varepsilon \geq 0$ which happens if and only if a firm's elasticity of demand is increase in their own-price.

E Strategic Complementarity under Kimball Demand

In the main text of the paper, I consider a generalization of the elasticities under CES aggregator and derive the strategic complementarities under this generalization. An alternative approach in the literature is using Kimball aggregator, which is also a generalization of the CES aggregator. In this section, I derive the demand functions of firms given this aggregator and show that the strategic complementarity implied by these demand functions cannot satisfy all of the following properties simultaneously: (1) there is weak strategic complementarity in pricing ($0 \leq \alpha < 1$), (2) there is substantial strategic complementarity in the data ($\alpha = 0.8$) and (3) strategic complementarity is decreasing with the number of firms within industries ($\frac{\partial \alpha}{\partial K} \leq 0$).

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

$$1 = K^{-1} \sum_{k \in K} f\left(\frac{KC_{j,k,t}}{\Phi(C_{j,1,t}, \dots, C_{j,K,t})}\right), \quad (31)$$

where $f(\cdot)$ is at least thrice differentiable, and $f(1) = 1$ (so that $\Phi(1, \dots, 1) = K$). 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 $P_{j,k,t} = J^{-1} Q_t \frac{\partial C_{j,k,t}}{\partial C_{j,t}}$, $\forall j, k$ where $C_{j,t} = \Phi(C_{j,1,t}, \dots, C_{j,K,t})$. Implicit differentiation of Equation (31)

gives

$$P_{j,k,t} = J^{-1}Q_t \frac{f'(\frac{KC_{j,k,t}}{C_{j,t}})}{\sum_{l \in K} C_{j,l,t} f'(\frac{KC_{j,l,t}}{C_{j,t}})}, \forall j, k. \quad (32)$$

To invert these functions and get the demand for every firm in terms of their competitors' prices, guess that there exists a function $F : \mathbb{R}^K \rightarrow \mathbb{R}$ such that $\frac{\sum_{l \in K} C_{j,l,t} f'(\frac{KC_{j,l,t}}{C_{j,t}})}{J^{-1}Q_t} = F(P_{j,1,t}, \dots, P_{j,K,t})$. I verify this guess by plugging in this guess to Equation (32), which implies the function $F(\cdot)$ is implicitly defined by $1 = K^{-1} \sum_{k \in K} f(f'^{-1}(P_{j,l,t} F(P_{j,1,t}, \dots, P_{j,K,t})))$. Note that this is consistent with the guess and $F(\cdot)$ only depends on the vector of these prices. It is straight forward to show that $F(\cdot)$ 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. Similar to the main text we can write this as

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

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 straight forward to show that η is the elasticity of substitution between industry goods around a symmetric point. Moreover, the elasticity of demand for every firms 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 $K \rightarrow \infty$ 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 $\alpha = (1 - \eta^{-1})K^{-1}$.

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

F Proofs of Propositions for the Dynamic Model

Proof of Proposition 3.

The optimality of one signal at any given time follows directly from Lemma 1 in [Afrouzi and Yang \(2019\)](#). Here, I include an adaptation of that proof for the special case of $\beta = 0$ that builds on the result in Lemma (B.7) for the dynamic case. Many arguments in the proof are similar and are omitted to avoid repetition. At a given time t , let $(S_{j,k}^{t-1})_{(j,k) \in J \times K}$ denote the signals that all firms have received until time $t - 1$, and are born with at time t . In particular, for any j, k , $S_{j,k}^{t-1} = (\dots, S_{j,k,t-3}, S_{j,k,t-2}, S_{j,k,t-1})$, where $\forall \tau \geq 1$, $S_{j,k,t-\tau} \subset \mathcal{S}^{t-\tau}$. This implies that (1) $S_{j,k,t-\tau}$ only contains information that were available at time $t - \tau$, and therefore are available at time t , and (2) $S_{j,k,t-\tau}$ is available for all other firms in the economy in case they find it desirable to learn about it.

Given this initial signal structure, pick a strategy profile for all firms at time t : $\zeta_t = (S_{j,k,t} \subset \mathcal{S}^t, p_{j,k,t} : S_{j,k,t}^t \rightarrow \mathbb{R})_{(j,k) \in J \times K}$, where $S_{j,k,t}^t = (S_{j,k,t}^{t-1}, S_{j,k,t})$. First, similar to the static case, we can show that in any equilibrium strategy $p_{j,k,t}(S_{j,k,t}^t)$ is linear in the vector $S_{j,k,t}^t$. This result follows with an argument similar to Lemma (B.3). Given this, let $p_{j,k,t}(S_{j,k,t}^t) = \sum_{\tau=0}^{\infty} \delta_{j,k,t}^{\tau} S_{j,k,t-\tau}$ denote the pricing strategy for any $(j, k) \in J \times K$. This is without loss of generality because the equilibrium has to be among such strategies. Notice that due to linearity and definition of \mathcal{S}^t , $p_{j,k,t}(S_{j,k,t}^t) \in \mathcal{S}^t$, $\forall (j, k) \in J \times K$. Now, pick a particular firm j, k and let $\zeta_{-(j,k),t}$ denote the signals and pricing strategies that ζ_t implies for all other firms in the economy except for j, k . Similar to Subsection B.4 let $\theta_{j,k,t}(\zeta_{-(j,k),t}) \equiv (q, (p_{j,l,t}(S_{j,l}^t))_{l \neq k}, (p_{m,n,t}(S_{m,n}^t))_{m \neq j, n \in K})'$ be the augmented vector of the fundamental, the prices of other firms in j, k 's industry, and the prices of all other firms in the economy. Now, define $\mathbf{w} = (1 - \alpha, \underbrace{\frac{\alpha}{K-1}, \dots, \frac{\alpha}{K-1}}_{K-1 \text{ times}}, \underbrace{0, 0, 0, \dots, 0}_{(J-1) \times K \text{ times}})'$. For the remainder

of the proof fix the capacity of the firm at $\kappa \geq 0$. I will show that the result holds for any such κ and hence is true also under the optimal κ . Since $\beta = 0$, fixing the capacity at some $\kappa \geq 0$, firm j, k 's signal choice problem is

$$\begin{aligned} & \min_{S_{j,k,t} \subset \mathcal{S}^t} \text{var}(\mathbf{w}' \theta_{j,k,t}(\zeta_{-(j,k),t}) | S_{j,k,t}^t) \\ \text{s.t. } & \mathcal{I}(S_{j,k,t}, \theta_{j,k,t}(\zeta_{-(j,k),t}) | S_{j,k,t}^{t-1}) \leq \kappa. \end{aligned}$$

To show that a single signal solves this problem, suppose not, so that $S_{j,k,t}$ contains more than one signal. Then, we know that $p_{j,k,t}(S_{j,k,t}^t) = \mathbf{w}' \mathbb{E}[\theta_{j,k,t}(\zeta_{-(j,k),t}) | S_{j,k,t}^t]$. Notice that I am assuming signals are such that these expectations exist. If not, then the problem of the firm is not well-defined as the objective does not have a finite value. To get around this issue, for

now assume that the initial signal structure of the game is such that expectations and variances are finite. Since both $\theta_{j,k,t}(\zeta_{-(j,k),t})$ and $S_{j,k}^t$ are Gaussian, $p_{j,k,t}(S_{j,k}^t) = \sum \delta_{j,k,\tau}^t S_{j,k,t-\tau}$ by Kalman filtering. Here for any $S_{j,k,t-\tau}$ that is not a singleton, let $\delta_{j,k,\tau}^t$ be a vector of the appropriate size that is implied by Kalman filtering. Therefore, by definition of \mathcal{S}^t , $p_{j,k,t}(S_{j,k}^t) \in \mathcal{S}^t$, meaning that there is a signal in \mathcal{S}^t that directly tells firm j,k what their price would be under $S_{j,k}^t$ and $\zeta_{-(j,k),t}$. Let $\hat{S}_{j,k}^t \equiv (S_{j,k}^{t-1}, p_{j,k}(S_{j,k}^t))$ and observe that by definition of $p_{j,k,t}(S_{j,k}^t)$, $\text{var}(\mathbf{w}'\theta_{j,k,t}(\zeta_{-(j,k),t})|S_{j,k}^t) = \text{var}(\mathbf{w}'\theta_{j,k,t}(\zeta_{-(j,k),t})|\hat{S}_{j,k}^t)$. Therefore, we have found a single signal that implies the same loss for firm j,k under $S_{j,k}^t$. Now, we just need to show that it is feasible, which is straight forward from data processing inequality: since $p_{j,k,t}(S_{j,k}^t)$ is a function $S_{j,k}^t$, we have

$$\mathcal{I}(p_{j,k,t}(S_{j,k}^t), \theta_{j,k,t}(\zeta_{-(j,k),t})|S_{j,k}^{t-1}) \leq \mathcal{I}(S_{j,k,t}, \theta_{j,k,t}(\zeta_{-(j,k),t})|S_{j,k}^{t-1}) \leq \kappa.$$

which concludes the proof for sufficiency of one signal. Now, given $S_{j,k}^{t-1}$ and $\theta_{j,k,t}(\zeta_{-(j,k),t})$ let $\Sigma_{j,k,t|t-1} \equiv \text{var}(\theta_{j,k,t}(\zeta_{-(j,k),t})|S_{j,k}^{t-1})$. Without loss of generality assume $\Sigma_{j,k,t|t-1}$ is invertible. If not, then there are elements in $\theta_{j,k,t}(\zeta_{-(j,k),t})$ 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 $\theta_{j,k,t}(\zeta_{-(j,k),t})$ to its orthogonal elements without limiting the signal choice of the agent. Now, for any non-zero singleton $S_{j,k,t} \in \mathcal{S}^t$, it is straight forward to show that

$$\mathcal{I}(S_{j,k,t}, \theta_{j,k,t}(\zeta_{-(j,k),t})|S_{j,k}^{t-1}) = \frac{1}{2} \log(1 - \mathbf{z}_t' \Sigma_{j,k,t|t-1}^{-1} \mathbf{z}_t),$$

where $\mathbf{z}_t \equiv \frac{\text{cov}(S_{j,k,t}, \theta_{j,k,t}(\zeta_{-(j,k),t})|S_{j,k}^{t-1})}{\sqrt{\text{var}(S_{j,k,t}|S_{j,k}^{t-1})}}$. The capacity constraint of the agent becomes $\mathbf{z}_t' \Sigma_{j,k,t|t-1}^{-1} \mathbf{z}_t \leq \lambda \equiv 1 - 2^{-2\kappa}$. Moreover, notice that the loss of the firm becomes

$$\text{var}(\mathbf{w}'\theta_{j,k,t}(\zeta_{-(j,k),t})|S_{j,k}^{t-1}, S_{j,k,t}) = \mathbf{w}' \Sigma_{j,k,t|t-1} \mathbf{w} - (\mathbf{w}' \mathbf{z}_t)^2.$$

This means that the agent can directly choose \mathbf{z}_t as long as there is a signal in \mathcal{S}^t that induces that covariance. I first characterize the \mathbf{z}_t that solves this problem and then show that such a signal exists. Notice that minimizing the loss is equivalent to maximizing $(\mathbf{w}' \mathbf{z}_t)^2$. The firm's problem is $\max_{\mathbf{z}_t} (\mathbf{w}' \mathbf{z}_t)^2$ s.t. $\mathbf{z}_t' \Sigma_{j,k,t|t-1}^{-1} \mathbf{z}_t \leq \lambda$. By Cauchy-Schwarz inequality we know $(\mathbf{w}' \mathbf{z}_t)^2 \leq (\mathbf{w}' \Sigma_{j,k,t|t-1} \mathbf{w})(\mathbf{z}_t' \Sigma_{j,k,t|t-1}^{-1} \mathbf{z}_t) \leq \lambda \mathbf{w}' \Sigma_{j,k,t|t-1} \mathbf{w}$, where the second inequality follows from the capacity constraint. Observe that $\mathbf{z}_t^* = \sqrt{\frac{\lambda}{\mathbf{w}' \Sigma_{j,k,t|t-1} \mathbf{w}}} \Sigma_{j,k,t|t-1} \mathbf{w}$ achieves this upper-bar. The properties of the Cauchy-Schwarz inequality imply that this is the only vector that achieves this upper-bar. Hence, \mathbf{z}_t^* is the unique solution to the firm's prob-

lem.⁵¹ Now, I just need to show that a signal exists in \mathcal{S}^t that implies this \mathbf{z}_t^* . To see this, let $S_{j,k,t}^* = (1 - \alpha)q_t + \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l,t}(S_{j,l}^t) + e_{j,k,t}$. It is straight forward to show that these signals imply \mathbf{z}_t^* .

Proof of Proposition 4.

Independence of α_j from j follows from the symmetry in the number of competitors across industries. Moreover, in the stationary equilibrium, all variances are constant over time. Since the choice of capacity depends on the underlying parameters and these variances, then the capacity is also time-invariant. 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 means that if firms are not learning about them, the uncertainty of firms is growing linearly over time and gets arbitrarily large. This implies that the benefit of learning monetary policy shocks is growing linearly with time which contradicts the choice of $\kappa = 0$. Thus, in the stationary equilibrium $\kappa > 0$ so that the conditional variance of the ideal price is stationary.

Now, from the proof of Proposition 3 recall that in the equilibrium, for all $(j,k) \in J \times K$, $p_{j,k,t}(S_{j,k}^t) = \mathbf{w}'\mathbb{E}[\theta_{j,k,t}(\zeta_{-(j,k),t})|S_{j,k}^t]$ where $S_{j,k}^t = (S_{j,k}^{t-1}, S_{j,k,t})$ and $S_{j,k,t} = (1 - \alpha)q_t + \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l,t}(S_{j,l}^t) + e_{j,k,t}$. From Kalman filtering

$$\begin{aligned} \mathbf{w}'\mathbb{E}[\theta_{j,k,t}(\zeta_{-(j,k),t})|S_{j,k}^t] &= \mathbb{E}[\mathbf{w}'\theta_{j,k,t}(\zeta_{-(j,k),t})|S_{j,k}^{t-1}] \\ &+ \frac{\mathbf{w}'\text{cov}(S_{j,k,t}, \theta_{j,k,t}(\zeta_{-(j,k),t}))}{\text{var}(S_{j,k,t}|S_{j,k}^{t-1})} (S_{j,k,t} - \mathbb{E}[S_{j,k,t}|S_{j,k}^{t-1}]). \end{aligned}$$

Notice from the proof of Proposition 3 that $\frac{\mathbf{w}'\text{cov}(S_{j,k,t}, \theta_{j,k,t}(\zeta_{-(j,k),t}))}{\text{var}(S_{j,k,t}|S_{j,k}^{t-1})} = \frac{\lambda}{\mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w}} \mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w} = \lambda$. Thus, using $p_{j,k,t}$ as shorthand for $p_{j,k,t}(S_{j,k}^t)$, $p_{j,k,t} = (1 - \lambda)\mathbb{E}[S_{j,k,t}|S_{j,k}^{t-1}] + \lambda S_{j,k,t}$. Finally, notice that $p_{j,k,t-1} = \mathbb{E}[S_{j,k,t-1}|S_{j,k}^{t-1}]$. Subtract this from both sides of the above equation to get $\pi_{j,k,t} \equiv p_{j,k,t} - p_{j,k,t-1} = (1 - \lambda)\mathbb{E}[\Delta S_{j,k,t}|S_{j,k}^{t-1}] + \lambda(S_{j,k,t} - p_{j,k,t-1})$, 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 $\pi_{j,k,t} = \mathbb{E}[\Delta S_{j,k,t}|S_{j,k}^{t-1}] + \frac{\lambda}{1-\lambda}(S_{j,k,t} - p_{j,k,t})$. Averaging this equation over all firms gives us the Phillips curve. To derive it, I take the average of every term separately and then sum them up.

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

⁵¹This solution can also be obtained by applying the Kuhn-Tucker conditions.

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 is approximately zero because J is large and $e_{j,k,t} \perp p_{m,l,t}, \forall m \neq j$, meaning that errors are orthogonal across industries regardless of coordination within them. 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 $\frac{\lambda}{1 - \lambda} = \frac{1 - 2^{-2\kappa}}{2^{-2\kappa}} = 2^{2\kappa} - 1$.

G Symmetric Stationary Equilibrium and Solution Method.

To characterize the equilibrium, I will use decomposition of firms' prices to their correlated parts with the fundamental shocks and mistakes as defined in the main text. I start with the fundamental q_t itself. Notice that since q_t has a unit root and is Gaussian, it can be decomposed to its random walk components: $q_t = \sum_{n=0}^{\infty} \psi_q^n \tilde{u}_{t-n}$, where $\tilde{u}_{t-n} = \sum_{\tau=0}^{\infty} u_{t-n-\tau}$, and $(\psi_q^n)_{n=0}^{\infty}$ is a summable sequence as Δq_t is stationary and $\Delta q_t = \sum_{n=0}^{\infty} \psi_q^n u_{t-n}$. In the case of $\beta = 0$, following Proposition 3 we know that given an initial signal structure for the game $(S_{j,k}^{-1})_{(j,k) \in J \times K}$, the equilibrium signals and pricing strategies are

$$\begin{aligned} S_{j,k,t} &= (1 - \alpha)q_t + \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,k,t}(S_{j,k}^t) + e_{j,k,t}, \\ p_{j,k,t}(S_{j,k}^t) &= \mathbb{E}[(1 - \alpha)q_t + \alpha \frac{1}{K-1} \sum_{l \neq k} p_{j,l,t}(S_{j,l}^t) | S_{j,k}^t] \\ &= \sum_{\tau=0}^{\infty} \delta_{j,k,t}^{\tau} S_{j,k,t-\tau}, \forall (j,k) \in J \times K, t \forall t \geq 0. \end{aligned}$$

However, these signals take a slightly different form when $\beta > 0$. Firms still receive one signal every period but they put different weights on the shocks. In particular, they realize that shocks to q are more persistent and taking the continuation value of knowing these

shocks into account focus more of their attention on the aggregate shocks. To solve for the optimal form of the signal, we follow the solution method outlined in [Afrouzi and Yang \(2019\)](#) for deriving the optimal weights on each of these shocks.

To characterize the equilibrium, given the form for the optimal signal, I do a similar decomposition analogous to the one in the static model. Given the pricing strategies of firms at time t , decompose their price to its correlated parts with the fundamental and parts that are orthogonal to it over time: $p_{j,k,t}(S_{j,k}^t) = \sum_{n=0}^{\infty} (a_{j,k,t}^n \tilde{u}_{t-n} + b_{j,k,t}^n v_{j,k,t-n})$. Here, $\sum_{n=0}^{\infty} b_{j,k,t}^n v_{j,k,t-n}$ is the part of j, k 's price at time t that is orthogonal to all these random walk components (mistake of firm j, k at time t). Moreover, $v_{j,k,t-n}$ is the innovation to j, k 's price at time t that was drawn at time $t - n$. In other words, I have also decomposed the mistake of the firm over time. This decomposition is necessary because other firms follow all these mistakes, but they can only do so after it was drawn at a certain point in time, in the sense that no firm can pay attention to future mistakes of their competitors as they have not been made yet. Before proceeding with characterization, I define the stationary symmetric equilibrium.

Definition 5. Given an initial information structure $(S_{j,k}^{-1})_{(j,k) \in J \times K}$, suppose a strategy profile $(S_{j,k,t} \in \mathcal{S}^t, p_{j,k,t} : S_{j,k}^t \rightarrow \mathbb{R})_{k \in K, t \geq 0}$ is an equilibrium for the game. We call this a symmetric steady state equilibrium if the pricing strategies of firms is independent of time, $t \geq 0$, and identity, $k \in K$. Formally, $\exists \{(a^n)_{n=0}^{\infty}, (b^n)_{n=0}^{\infty}\}$, such that $\forall t \geq 0, \forall (j, k) \in J \times K$, $p_{j,k,t} = \sum_{n=0}^{\infty} (a^n \tilde{u}_{t-n} + b^n v_{j,k,t-n})$.

To characterize the equilibrium, notice that we not only need to find the sequences $(a^n, b^n)_{n=0}^{\infty}$, but also the joint distribution of $v_{j,k,t-n}$'s across the industries. To see this, take firm j, k and suppose all other firms are setting their prices according to $p_{j,k,t} = \sum_{n=0}^{\infty} (a^n \tilde{u}_{t-n} + b^n v_{j,k,t-n})$. Then, firm j, k 's optimal signals are given by the solution method outlined in [Afrouzi and Yang \(2019\)](#). In the special case of $\beta = 0$ this form is given by

$$S_{j,k,t} = \sum_{n=0}^{\infty} \left[((1 - \alpha)\psi_q^n + \alpha a^n) \tilde{u}_{t-n} + \alpha b^n \frac{1}{K-1} \sum_{l \neq k} v_{j,l,t-n} + e_{j,k,t} \right],$$

where by properties of the equilibrium $e_{j,k,t}$ is the rational inattention error and is orthogonal to \tilde{u}_{t-n} and $v_{j,l,t-n}$, $\forall n \geq 0, \forall l \neq k$. Using the joint distributions of errors $(v_{j,k,t-n})_{k \in K}$, by Kalman filtering, the firm would choose to set their price according to

$$p_{j,k,t} = \sum_{n=0}^{\infty} \delta^n S_{j,k,t-n} = \sum_{n=0}^{\infty} (\tilde{a}_n \tilde{u}_{t-n} + \tilde{b}_n \frac{1}{K-1} \sum_{l \neq k} v_{j,k,t-n} + \tilde{c}_n e_{j,k,t-n})$$

for some sequences $(\tilde{a}_n, \tilde{b}_n, \tilde{c}_n)$. But in the equilibrium, $p_{j,k,t} = \sum_{n=0}^{\infty} (a^n \tilde{u}_{t-n} + b^n v_{j,k,t-n})$. This implies, $a^n = \tilde{a}^n$, $b^n v_{j,k,t-n} = \tilde{b}_n \frac{1}{K-1} \sum_{l \neq k} v_{j,l,t-n} + \tilde{c}_n e_{j,k,t-n}$, where $e_{j,k,t-n} \perp v_{j,l,t-n}, \forall l \neq k$

k. Using the second equation we can characterize the joint distribution of $(v_{j,k,t-n})_{k \in K}, \forall n \geq 0$. This joint distribution is itself a fixed point and should be consistent with the Kalman filtering behavior of the firm that gave us $(\tilde{a}_n, \tilde{b}_n, \tilde{c}_n)_{n=0}^{\infty}$ in the first place. Finally, notice that underneath all these expressions we assume that these processes are stationary meaning that the tails of all these sequences should go to zero. Otherwise, the problems of the firms are not well-defined and do not converge. I verify this computationally, by truncating all these sequences such that $\forall n \geq \bar{T} \in \mathbb{N}, a^n = b^n = 0$ where \bar{T} is large, solving the problem computationally, and checking whether the sequences go to zero up to a computational tolerance before reaching \bar{T} . In my code I set $\bar{T} = 100$. The economic interpretation for this truncation is that all real effects of monetary policy should disappear within 100 quarters. Such truncations are the standard approach in the literature for solving dynamic imperfect information models.

The following algorithm illustrates my method for solving the problem.

Algorithm 1. *Characterizing a symmetric stationary equilibrium:*

1. Start with an initial guess for $(a^n, b^n)_{n=0}^{\bar{T}-1}$, and solve for a representative firm j, k 's optimal signal using the method in [Afrouzi and Yang \(2019\)](#) (for the case of $\beta = 0$ set $S_{j,k,t} = \sum_{n=0}^{\bar{T}-1} \left[((1-\alpha)\psi_q^n + \alpha a^n) \tilde{u}_{t-n} + \alpha b^n \frac{1}{K-1} \sum_{l \neq k} v_{j,l,t-n} + e_{j,k,t} \right]$).
2. Using Kalman filtering, given the set of signals implied by previous step, form the best pricing response of a firm and truncate it. Formally, find coefficients $(\tilde{a}_n, \tilde{b}_n, \tilde{c}_n)_{n=0}^{\bar{T}-1}$ such that $p_{j,k,t} \approx \sum_{n=0}^{\bar{T}-1} (\tilde{a}_n \tilde{u}_{t-n} + \tilde{b}_n \frac{1}{K-1} \sum_{l \neq k} v_{j,l,t-n} + \tilde{c}_n e_{k,t-n})$.
3. $\forall n \in \{0, \dots, \bar{T}-1\}$, update $a^n = \tilde{a}^n$, and b^n such that $b^n v_{k,t-n} = \tilde{b}_n \frac{1}{K-1} \sum_{l \neq k} v_{j,l,t-n} + \tilde{c}_n e_{k,t-n}$, using $e_{k,t} \perp v_{-k,t}$, and the symmetry of the distribution of $(v_{j,k,t})_{k \in K}$.
4. Iterate until convergence of the sequence $(a^n, b^n)_{n=0}^{\bar{T}-1}$.

H A Static Model with Heterogeneous Market Shares

Consider the household's demand with CES aggregator from Equation 7 with the following modification:

$$C_t = \prod_{j \in J} \left[\left(\sum_{k \in K_j} \tilde{m}_{j,k}^{\frac{1}{\eta}} C_{j,k,t}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \right]^{J-1} \quad (33)$$

where now $\tilde{m}_{j,k}$ captures the taste of the consumer for the product of firm k in industry j . Moreover, $\forall j$ we normalize $\sum_k \tilde{m}_{j,k} = 1$ 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

$$C_{j,k,t} = P_t C_t \frac{\bar{m}_{j,k} P_{j,k,t}^{-\eta}}{\sum_l \bar{m}_{j,l} P_{j,l,t}^{1-\eta}} \quad (34)$$

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

$$\varepsilon_{j,k,t} = \eta - (\eta - 1) \frac{\bar{m}_{j,k} P_{j,k,t}^{1-\eta}}{\sum_l \bar{m}_{j,l} P_{j,l,t}^{1-\eta}} \quad (35)$$

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 18, is given by

$$\max_{\{\kappa_{j,k,t}, S_{j,k,t}, p_{j,k,t}(S_{j,k}^t)\}_{t \geq 0}} -\mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t \left(\underbrace{\eta (p_{j,k,t}(S_{j,k}^t) - p_{j,k,t}^*)^2}_{\text{loss from mispricing}} + \underbrace{\omega \kappa_{j,k,t}}_{\text{cost of capacity}} \mid S_{j,k}^{-1} \right) \right] \quad (36)$$

$$\begin{aligned} \text{s.t. } p_{j,k,t}^* &\equiv (1 - \alpha_{j,k}) q_t - \alpha_{j,k} p_{j,-k,t}(S_{j,-k,t}) \\ \mathcal{I} \left(S_{j,k,t}, (q_\tau, p_{l,m,\tau}(S_{l,m}^\tau))_{0 \leq \tau \leq t}^{(l,m) \neq (j,k)} \right) &\leq \kappa_{j,k,t} \\ S_{j,k}^t &= S_{j,k}^{t-1} \cup S_{j,k,t}, \quad S_{j,k}^{-1} \text{ given.} \end{aligned} \quad (37)$$

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:

$$\alpha_{j,k} = (1 - \eta^{-1}) \bar{m}_{j,k} \quad (38)$$

Note that in this case $\bar{m}_{j,k}$ 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 $\eta \rightarrow \infty$.⁵² Then, taking

⁵²In this hypothetical example, having $\eta \rightarrow \infty$ 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

a second-order approximation around this steady state, it follows from Equation 15 that the ideal price of firm j, k is given by

$$p_{j,k,t}^* = (1 - \bar{m}_{j,k})q_t + \bar{m}_{j,k} \frac{\sum_{l \neq k} \bar{m}_{j,l} p_{j,l,t}}{\sum_{l \neq k} \bar{m}_{j,l}} \quad (39)$$

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

almost perfect information. To resolve this, we assume that ω is also proportional to η so that the ratio stays constant as $\eta \rightarrow \infty$.