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

Hassan Afrouzi†

October 12, 2017

Abstract

In countries where inflation has been low and stable, price setters display highly dispersed aggregate inflation expectations; especially so when they face fewer competitors. In contrast to the predictions of standard models, realized inflation deviates significantly from price setters' aggregate inflation expectations. Instead, their own-industry inflation expectations are more accurate, and aggregate inflation tracks these expectations closely. I propose a new dynamic model of rational inattention with oligopolistic competition to explain these stylized facts. The Phillips curve relates aggregate inflation to price setters' own-industry inflation expectation, and firms forego learning about aggregate variables to focus on their own-industry prices. This incentive is stronger when every firm faces fewer competitors. Using new firm-level survey evidence, I calibrate the degree of rational inattention and industry size in the model and find that a two-fold increase in the number of competitors reduces the half-life and on-impact response of output to a monetary policy shock by 40 and 15 percent, respectively.

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

JEL classification: D21; D43; D83; E31; E32

*This paper is a revised version of the first chapter of my PhD dissertation at UT Austin. I would like to thank my advisors Olivier Coibion, Saroj Bhattarai and Yuriy Gorodnichenko for their guidance and support. I also benefited from discussions with V. Bhaskar, Nicholas Buchholz, Mark Dean, Andrew Glover, Matthias Kehrig, Filip Matejka, Eugenio Miravete, Ricardo Reis, Christopher Sims, Max Stinchcombe, Mirko Wiederholt, Emily Weisburst, Michael Woodford and numerous seminar participants. All errors are my own.

†Columbia University. Email: hassan.afrouzi@columbia.edu.

"[I]t is far, far safer to be wrong with the majority than to be right alone."

John Kenneth Galbraith (1989)

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] This insight has profoundly influenced monetary policy: central bankers treat anchored expectations not only as a policy objective for controlling inflation, but also as a potential instrument since the onset of the zero lower bound after the Great Recession.

In spite of this consensus, empirical evidence on price setters' expectations of aggregate inflation are at odds with this theoretical prediction. For instance, Kumar, Afrouzi, Coibion, and Gorodnichenko (2015) show that despite a long history of low and stable inflation in New Zealand, firms in that country are exceedingly uninformed about it.[2] Managers make average errors of 2 to 3 percentage points in perceiving current as well as forecasting future inflation.[3] They revise their forecasts by an average of 3.4% after only three months, and report an average standard deviation of 1% around their inflation forecasts. 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. While the theory predicts that such high and volatile expectations of aggregate inflation should either pass through to inflation or be accompanied by a deep contemporaneous recession, neither was the case for these countries at the time of the surveys.

One cannot reconcile these empirical observations with our current models. Either the survey data are inaccurate in reflecting expectations, or our baseline models are too simplified to capture the channels that would explain this disconnect. With respect to the latter, I show in this paper how stable inflation is consistent with such volatile expectations by introducing a new dynamic model of rational inattention with oligopolistic competition. Firms optimally, and strategically, choose to learn more about the prices of their competitors at the expense of knowing less about aggregate variables. I calibrate the degree of rational inattention and

[1] The timing of these expectations are model-specific. For instance, New Keynesian sticky price models relate inflation to expectations of future aggregate inflation, while imperfect information models, pioneered by Lucas (1972), relate it to past expectations of current inflation.

[2] With a slight abuse of terminology, throughout the paper, I refer to a firm's managers' expectations as firm's expectations.

[3] A manager's perception of inflation is defined as their nowcast of current inflation. In other words, this perception is the expectation that is formed over current inflation.

the number of firms within industries using firm-level survey evidence and show that these features have significant macroeconomic implications for the propagation of monetary policy shocks to inflation and output.

Although rational inattention is not the only explanation for the large and persistent forecast errors observed in the expectations of managers, complementary evidence from countries with volatile inflation is inconsistent with alternative explanations such as financial illiteracy, scarcity of information, and either complexity or lack of transparency in monetary policy. Comparing the U.S. and Argentina, Cavallo, Cruces, and Perez-Truglia (2014) show that individuals in lower inflation contexts have significantly weaker priors about the inflation rate, a finding that supports the rational inattention hypothesis. Moreover, in a recent and ongoing project with the Central Bank of Iran (CBI), I conduct a survey of firms' expectations in Iran, a country that has been dealing with highly volatile inflation over the last four decades. Despite the fact that inflation has ranged from 9% to 40% over the last three years, firms' inflation expectations are relatively precise. Their average expectation is only 2 percentage points away from the realized inflation, and despite the high volatility of inflation, the dispersion of their expectations is only 3.5%. This evidence not only corroborates the rational inattention narrative but also casts doubt on the aforementioned alternative explanations, as it is highly unlikely that households or firm managers in developed countries are less literate or have less access to information about monetary policy than managers in Iran or households in Argentina.

The other building block of the model is the role of imperfect competition at the micro-level. Although the economy consists of a large number of firms, each one of them only competes directly with a finite number of others at the micro-level. When asked how many competitors they face in their main product market, firms in New Zealand report only between 5 to 8 rivals on average, with 35% of firms responding that they face fewer than 4 competitors, and only 5% reporting that they have more than 15 competitors.

Firms that compete with only a few others do not optimize over their price relative to an aggregate price index, as is implied by standard models in which a firm is one of a continuum, but rather relative to the prices of their direct rivals, a feature which has important implications for monetary policy when information acquisition is endogenous. Every firm realizes that their rationally inattentive competitors will make mistakes in perceiving the shocks to the economy, which could hurt the firm by pushing it away from its optimal price.[4] Therefore, even in an economy with a single aggregate shock, firms find themselves facing an endogenous trade-off: how much to track the shock itself versus the mistakes of others.

[4] I define what I precisely mean by "mistakes" in the main body of the paper. In short, a mistake is the part of a firm's price which is unpredictable by the fundamental shocks of the economy.

In other words, a firm's problem is to find an optimal middle ground between *being more wrong about the fundamental shocks with the majority of their competitors* and *having more precise estimates of those fundamentals by themselves*.

The first contribution of this paper is to show that in the presence of this trade-off, inflation will depend not just on price setters' expectations of aggregate inflation but also on their expectations of own-industry inflation. Furthermore, when there is enough strategic complementarity in price setting within industries, inflation dynamics will depend primarily on the latter and little on the former. This can therefore account for how, 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. The latter simply play little role in price setting decisions when rationally inattentive price setters have strategic motives, which drive them to have more precise information about the prices of their competitors at the cost of being less informed about aggregate inflation.

The second contribution of this paper is to characterize the incentives of firms in tracking the mistakes of their competitors, and document the quantitative implications of these incentives in the propagation and amplification of monetary policy shocks in a dynamic general equilibrium model. Within economies where firms face more direct competitors, they have lower incentives in tracking the mistakes of their rivals because it is more unlikely for a larger group of competitors to make a mistake on average. Consequently, firms facing a larger number of competitors allocate a higher amount of their attention to learning the monetary policy shocks. Therefore, it takes a shorter time for these firms to fully realize the magnitude of a shock to monetary policy and adjust their prices accordingly, which in turn translates to a lower persistence in the real effects of a monetary policy shock. I show that a two-fold increase in the number of competitors that every firm faces directly at the micro-level of the economy decreases the half-life of output and inflation responses to a one percent monetary policy shock by 40 and 25 percent, while reducing their on-impact response by 15 and 33 percent respectively.

Another contribution of this paper is calibrating the capacity of processing information, which has been a difficult task for rational inattention literature so far due to lack of suitable data. Newly available data on price setters' expectations from New Zealand, however, creates an ideal ground for the calibration of this parameter by directly measuring the degree of information rigidity in firms' forecasts of aggregate inflation. Moreover, by deriving the Phillips curve within the rational inattention model, I also relate the capacity of processing information to its analogous parameters in other models of information rigidity, namely noisy and sticky information models such as Woodford (2003a) and Mankiw and Reis (2002) respectively, as well as empirical literature that estimates these rigidities using survey data, namely Coibion and Gorodnichenko (2012, 2015). I show that this capacity directly maps to the Kalman gain, the weight that firms put on their new information, in noisy information models, and the measure of firms that update their information within a sticky information model, allowing for a simple comparison of the degree of information rigidity across models. Utilizing this relationship, I find that the calibrated value of capacity for processing information implies a much lower degree of information rigidity than commonly needed in noisy and sticky information models. price setters in my model ultimately are very good at processing information, but spend a portion of that attention to tracking the mistakes of their competitors due to their strategic incentives rather than tracking macroeconomic variables. Hence, in spite of being well-informed about their own optimal prices, price setters' macroeconomic beliefs endogenously become akin to those of agents facing large information rigidities for macroeconomic variables.

The theoretical approach of this paper is closely related to the literature on endogenous information acquisition in beauty contests.[5] In their seminal paper, Hellwig and Veldkamp (2009) formalize the idea of tracking others' beliefs in a setting with a measure of agents and show that the value of public information increases with strategic complementarity in actions.[6] Within this literature, the closest paper to this one is Denti (2015) who formalizes static information acquisition games with a countable number of players and an unrestricted information structure, and shows that in such large games, players' signals are independent conditional on the fundamental. I also consider a large game by modeling the production side of the economy, but focus on a case where every firm directly competes with a few others. Therefore, in spite of having a large number of firms in the model, they optimally choose to track the mistakes of their few competitors, and directly track their beliefs. Also, a major departure of my paper from this literature is that it is the first one to investigate the implications of these oligopolistic incentives in a dynamic stochastic general equilibrium model, and show that firms' incentives in tracking one anothers' beliefs can have important implications for the magnitude as well as the persistence of output and inflation in response to monetary policy shocks.

Finally, this paper also builds on the rational inattention literature and the seminal work of Sims (2003). Maćkowiak and Wiederholt (2009, 2015) show how rational inattention on the part of firms and households affect the dynamics of inflation and output in the economy. While this literature has assumed that firms' signals are independent conditional on the fundamental shocks, I mainly depart from this literature by micro-founding the endogenous

[5]For a comprehensive recent survey of this literature see Angeletos and Lian (2016).

[6]In a similar setting, Myatt and Wallace (2012) show that the endogenous information acquisition of agents becomes more public in nature as the degree of strategic complementarity increases. Colombo et al. (2014) show how the acquisition of private information affects the value of public information.

strategic interactions of agents in tracking mistakes of one another and show that this assumption holds only if firms compete with a measure of others. When, instead, the number of competitors for every firm in the economy is finite, firms choose signals that incorporate correlated errors. The dynamic model of this paper also relates to a very recent literature on characterizing dynamic incentives in information acquisition. Mackowiak, Matejka, and Wiederholt (2016) show that rational inattention leads to a forward looking behavior in information acquisition of agents. Furthermore, by formalizing the dynamic incentives of agents in acquiring information, Afrouzi and Yang (2016) show that agents' optimal information acquisition strategy in dynamics is based on motives of information smoothing over time. This paper departs from this literature by focusing mainly on strategic incentives of firms rather than their dynamic incentives.

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. Section 5 discusses the impulse responses of the calibrated model. Section 6 concludes. Moreover, all the technical derivations as well as the proofs of all the propositions and corollaries are included in Appendices A and B, for the static and dynamic models respectively.

2 A Static Model

The goal of this section is to endogenize informational choices of imperfectly competitive firms and illustrate the equilibrium relationship between aggregate price and the expectations of firms within a static model. The model presented here is a special case of the dynamic general equilibrium model that is specified in Section 4. While the general dynamic model has to be solved using computational methods, the solution to the static case is in closed form, which provides insight for interpreting the results of the dynamic model.

Since the main purpose of this section is to provide intuition, I focus on the economics of the forces at work in the main text. All informal claims in this section are formalized in Appendix A, and the proofs for propositions are included in Appendix A.8.

2.1 The Environment

There are a large number of industries in the economy indexed by $j \in \{1, \ldots, J\}$, and within every industry there are K firms. Let index j, k denote firm k in industry j. Firms are price setters and pay attention to a normally distributed fundamental that I denote by

$q \sim \mathcal{N}(0,1)$, which represents the underlying source of randomness in the objectives of these firms. I interpret q as the nominal aggregate demand in this economy, and assume that it is exogenously drawn by a central bank.

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 the distance between their price and a convex combination of q and the average of their competitors' prices;

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

where $\alpha \in [0,1)$ denotes the degree of strategic complementarity within industries.[7] Two assumptions in the specification of this environment are essential for the results that will follow, which are also the main departures of this paper from the baseline models of inflation dynamics with imperfect information. The first assumption is the existence of strategic complementarity *within industries*, and the second is the finiteness of the number of competitors within them.

To illustrate the implications of imperfect information in this environment, it is useful to consider a reduced form example before we move on to the rational inattention problem of the firms. For an endowed information set for the economy, let $\mathbb{E}^{j,k}[.]$ be the expectation operator of firm j,k. Every firm chooses the price that minimizes its expected loss:

$$\forall j,k,\ p_{j,k} = \text{argmin}_x \mathbb{E}^{j,k}[(x - (1-\alpha)q - \alpha p_{j,-k})^2],$$

where $p_{j,-k}$ denotes the average price of firm k's competitors.[8] Aggregating the best response of the firms across the economy, we get the following expression for the aggregate price,

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

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 the assumption on finiteness of firms within industries. The aggregate price no longer depends on the average expectation of the aggregate price across firms, but the average expectation of their own-industry prices. In fact, when α is large, as the data will strongly suggest in Section 3, it is mainly the latter

[7] Here 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.

[8] See, for instance, Morris and Shin (2002); Angeletos and Pavan (2007) for a discussion of such games with exogenous information sets, and the value of information within them.

that drives the aggregate price.

Although a static model inherently does not incorporate the concept of inflation, in this case, the intuition that the static model gives us, regarding aggregate prices and expectations of firms, carries on to the Phillips curve of the dynamic economy. In short, similar to the aggregate price in the static model, aggregate inflation in the dynamic model is mainly driven by firms' average expectations of *own-industry price changes* rather than by their expectations of aggregate inflation.

Therefore, in order to understand how prices are determined in the economy, we need to understand how firms form their expectations of both the fundamental as well as the prices of their competitors.

2.2 The Information Choice Problem of the Firms

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

I model the information choice problem of the firms following the rational inattention literature. The spirit of rational inattention is the richness of *available* information that it assumes for an economy. This in itself separates a rational inattention economy from one with an information structure in which agents either observe a set of exogenously imposed signals or choose their signals from a set that does not allow for sufficiently precise signals. In a rational inattention world, however, if an action takes place after the nature draws a random source, then perfect information of that source is available for the agents. For instance, if firms are setting their prices after a monetary policy shock has taken place, it is unreasonable to assume that they do not have access to its exact realization, which is also the primary building block of the full-information rational expectations hypothesis. What distinguishes rational inattention from full information rational expectations, however, is the recognition of the fact that *availability* of information is a different notion that its *feasibility* for the firms. The fact that perfect information is available about a monetary policy shock does not necessarily imply that firms would *choose* to have perfect information when attention is costly. Nonetheless, subject to this cost, firms behave optimally and choose their information set such that it maximizes their ex ante payoffs.

Appendix A.4 shows that if the set of available signals $\mathcal{S}$ is *rich* enough, rationally inattentive firms always prefer to observe a single signal, rather than observing multiple ones.[9] The intuition behind this result is that if there is enough variation in the sources of news within the economy, there is always a single signal available that is precisely what the manager would like to see, subject to their limited attention.[10]

Therefore, a pure strategy for any firm j,k is to choose a signal, $S_{j,k} \in \mathcal{S}$, and a pricing strategy that maps the realization of the signal into the firm's price, $p_{j,k} : S_{j,k} \to \mathbb{R}$. I show in Appendix A.3 that in any equilibrium pricing strategies are linear in firms' signals. I take this result as given here and focus on linear strategies, where firm j,k chooses $M_{j,k} \in \mathbb{R}$, such that $p_{j,k} = M_{j,k} S_{j,k}$. Given a strategy profile for all other firms in the economy, $(S_{l,m}, M_{l,m})_{(l,m)\neq(j,k)}$, firm j,k's rational inattention problem is

$$\begin{aligned} &\min_{(S_{j,k}\in\mathcal{S}, p_{j,k}:S_{j,k}\to\mathbb{R})} \mathbb{E}[(p_{j,k} - (1-\alpha)q - \alpha \tfrac{1}{K-1}\textstyle\sum_{l\neq k} M_{j,l}S_{j,l})^2 | S_{j,k}] \\ s.t. \quad &\mathcal{I}(S_{j,k}; (q, M_{l,m}S_{l,m})_{(l,m)\neq(j,k)}) \leq \kappa \end{aligned} \tag{2}$$

where $\mathcal{I}(S_{j,k}; (q, M_{l,m}S_{l,m})_{(l,m)\neq(j,k)})$ measures the amount of information that the firm's signal reveals about the fundamental and the prices of other firms in bits.[11] 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 $\mathcal{S}$. The following defines an equilibrium for this economy.

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

It is shown in Appendix A that while there are many equilibria for this game in terms of the signals that firms choose, all of them are equivalent and unique in one sense: all equilibria point toward a unique joint distribution in prices of firms. This implies that while many equilibria could be formed regarding what sources of news firms follow, the joint distribution of prices is unique and remains unchanged and independent of which equilibrium is being played by the firms.

[9] See Section A.2 for a formal definition of a rich information structure. My definition of a rich information set corresponds to the concept of flexibility in information acquisition in Denti (2015).

[10] To see this, suppose a manager's optimal strategy is to see two distinct Gaussian signals, that are not perfectly correlated, and maps them into their firms' price using a linear strategy. But notice that this price cannot be a sufficient statistic of the underlying signals as its dimension is strictly smaller. This means there is information in manager's information set that is not used by them in pricing. Therefore, the manager would prefer to deviate to a signal that eliminates the information that they do not use but is more precise in terms of the information that they do, which is feasible if the information structure is rich enough to allow for such a deviation. This is a contradiction with optimality of two distinct signals. Hence, the optimal signal structure has to be one-dimensional.

[11] $\mathcal{I}(.;.)$ is Shannon's mutual information function. In this paper, I focus on Gaussian random variables, in which case $\mathcal{I}(X;Y) = \frac{1}{2}\log_2(\det(var(X))) - \frac{1}{2}\log_2(\det(var(X|Y)))$.

To discuss the incentives that form the equilibrium, I introduce the following reinterpretation of a firm's problem. The uniqueness of the joint distribution of prices in the equilibrium allows us to abstract from 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. The finite attention of the firm implies that this price cannot be fully revealing of the fundamental, as figuring out the fundamental with infinite precision requires infinite attention on the part of the firm. Therefore, any firm's price can be decomposed into a part that is correlated with the fundamental and a part that is orthogonal to it:

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

The vector $(v_{j,k})_{j,k \in J \times K}$ contains the *mistakes* of all firms in pricing, with their joint distribution being endogenously determined in the equilibrium.[12] I define these orthogonal elements mistakes because in a world where firms have infinite capacity to process information, all firms perfectly learn the fundamental and set their prices exactly equal to q. Since this concept is going to be an important part of my argument in this paper, the following definition formally characterizes it.

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

This sheds light on the economic importance of this decomposition. A mistake is a source of volatility in a firm's price that is unpredictable by anyone who only knows the fundamentals of the economy, and nothing more about the firm's price. It is important to mention that these mistakes need not to be independent across firms. In fact, by endogenizing the informational 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 learn about the mistakes of their rivals and incorporate them in their own decisions.

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. For notational ease, let $\lambda \equiv 1 - 2^{-2\kappa} \in [0, 1)$ be a transformation of a firm's capacity for processing information. $\lambda = 0$ corresponds to a complete absence of the ability to process information, $\kappa = 0$, and $\lambda \to 1$ corresponds to an infinite capacity to do so, $\kappa \to \infty$. From

[12] Mistakes need not to be independent. In fact, it is at the heart of this paper to understand how these mistakes are jointly distributed.

here onwards, I denote λ as the total amount of *attention* that a firm has at its disposal. The following definition formally specifies this term.

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, for any two random variables X and Y, 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.[13] Appendix A 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:

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

Here $\sigma_v \equiv var(v_{j,-k})^{\frac{1}{2}}$ 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 information processing constraint reduces such that the square of the two correlations should sum up to an amount less than λ. Figure (1) symbolically illustrates the feasible set of correlations and the indifference curve for the values that maximize the problem above.

The following proposition states the properties of the equilibrium. The closed form solutions and derivations are included in Appendix A. I focus here on the forces that shape this equilibrium.

Proposition 1. *In equilibrium,*

1. *Firms pay attention not only to the fundamental, but also to the mistakes of their competitors:* $\rho_v^* > 0$.

2. *A firms' knowledge of the fundamental increases in the number of their competitors and decreases in the degree of strategic complementarity:*

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

[13] For 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 $\rho_{X,Y}^2$.

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 independence of mistakes from the fundamental implies an endogenously arisen trade-off for firms in allocating their attention. Higher attention to competitors' mistakes *has* to be compensated by lower attention to the fundamental but this in turn reduces a firm's losses by creating coordination between them and their rivals. The existence of this trade-off is founded on a firm knowing that their competitors make mistakes. In a world where a firm's competitors never made mistakes, all the firm would need to know was the fundamental. This has an important implication regarding what firms know about the fundamental in the equilibrium: price setters who pay more attention to the mistakes of others have less information about the fundamental.

The presence of σ_v in the objective of the firm unveils another force in determining the incentive of a firm in paying attention to others' mistakes. The firm cares not about the mistake of any single competitor, but about the average mistake that its rivals make all together. If σ_v is zero, then the firm does not have any incentive to pay attention to any possible mistakes made by any individual rival. Therefore, the more the mistakes of a firm's competitors "wash out", the less the firm is worried about them. Intuitively, when the average mistakes wash out, the firm is fully confident that while there are some rivals that undercut them by mistake, there are others that compensate by overpricing them. In other words, a firm's profits, on average, will not be affected by the mistakes of other firms. Formally, this dependence to the number of competitors can be seen in the expression for the variance of others' mistakes when they play the equilibrium strategy:

$$\sigma_v^2 = \sigma^2 \left(\frac{1}{K-1} + \frac{K-2}{K-1}\rho \right),$$

where ρ is the correlation of any two competitors' mistakes and σ^2 is their variance, both of which are firm independent due to the symmetry of the equilibrium strategy. As $\rho < 1$ due to finite capacity, σ_v^2 is strictly decreasing in K.[14] Thus, the firm's incentive in paying attention to the mistakes of its competitors diminishes as their number increases. In fact, in the equilibrium as $K \to \infty$, σ_v^2 converges to zero, and firms lose all their interest in learning about the mistakes of their rivals. In other words, large industries never find it optimal to coordinate once the information structure is rich enough to allow them to choose their

[14]This is true for a given σ^2, as well as taking into account how σ^2 would endogenously change as a result of changing K. Moreover, the claim that $\rho < 1$ holds due to the fact that firms can never perfectly correlate their mistakes if they have finite attention span. Perfect coordination requires infinite precision in acquiring information and therefore infinite capacity to process information.

optimal degree of coordination: if industries are large enough, given any degree of positive correlation in the mistakes of others, the law of large numbers washes out the uncorrelated part of those mistakes.[15] Since the firm only cares about those mistakes by a coefficient $\alpha < 1$, it would find it optimal to choose a degree of coordination that is less than what others have chosen. This implies that in the limit as $K \to \infty$ where the law of large numbers holds for industries, the only plausible equilibrium is when there is no coordination. If everybody else chooses not to coordinate, the mistakes wash out, and for any single firm, there is no average mistake to track.

Another important aspect of the Proposition 1 is how strategic complementarity influences the choices of these firms. α is the underlying parameter that relates the payoff of a firm to mistakes of its competitors. When α is zero, the firm pays no penalty for charging a price that is farther away from the prices of its competitors, implying that the firm's payoff depends only on how close its price is to the fundamental itself. Since tracking the mistakes of others is costly in terms of learning the fundamental, when $\alpha = 0$, all firms focus solely on the fundamental and learn about it as much as their finite attention allows them. As α gets larger, however, the payoffs of firms depend more on the mistakes that others make and 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.

Appendix A shows that in equilibrium

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

This implies that the degree to which prices covary with the fundamental in an industry depends on strategic complementarity and the capacity of processing information while it is independent of the number of firms in the industry. We will show this independence is specific to the static environment and goes away in the micro-founded dynamic model. However, even within this static environment, this result holds a key insight into the relationship of expectations and prices.[16]

Moreover, the fact that higher strategic complementarity decreases δ is a well-known result in games of incomplete information. Higher capacity for processing information, on

[15] This is analogous to the result in Denti (2015) which shows that within large games coordination vanishes once the information structure is unrestricted.

[16] While Proposition 1 shows that the quality of firms' information about the fundamental is lower when K is smaller, the independence of the price level from K points toward the fact that this lower information quality about the fundamental is compensated by better information of firms about the prices of their competitors. In the static model, these two forces cancel each other out and make δ independent of K. I discuss the economic nature of this independence in more detail in the discussion that follows Proposition 7.

the other hand, increases this covariance: a firm with a higher ability to follow the news is able to respond more confidently to that news. In particular, the degree of strategic complementarity becomes irrelevant as firms approach infinite capacity to process information. This corresponds to an environment where firms can perfectly observe the fundamental and reach common knowledge about it. Once they do so, strategic complementarity is irrelevant to their action as everyone knows the fundamental perfectly, knows that its opponents know the fundamental perfectly, and so it goes in the hierarchy of the higher order beliefs. This common knowledge allows each firm to respond to their knowledge of the fundamental with the highest confidence and set their price equal to that, while absolutely minimizing their losses.

2.3 Equilibrium Prices and Expectations

Having characterized the equilibrium, we now have the necessary tools to answer our motivating question on the relationship between equilibrium prices and expectations. Recall from equation (1) that in the equilibrium the average price is given by

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

Here, the goal is to understand how the aggregate price co-moves with the average expectations of firms from the objects of the model. The next proposition derives the necessary results for the argument that follows.

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*

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

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

Therefore, what matters for the determination of the aggregate price is not what firms know about the aggregate price level, but what they know about their own industry prices. This result also holds in the dynamic model in the sense that inflation is driven more by the expectations of industry price changes, than 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:

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

where $u_j \perp p$ is the common mistake in industry j and $e_{j,k}$ is the independent mistake of firm j,k. The fact that $var(u_j) \neq 0$ by Proposition 1 implies that the firm would be more confident in predicting their own industry price changes than the aggregate price, and the two would become the same only if there was no coordination within industries, which happens when $K \to \infty$.

This result, along with its counterpart in the dynamic model, shows how stable inflation can be an equilibrium outcome even when agents' 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, aggregate price does not by itself play an important role in firms' profits so they do not need to directly learn about it.

Thus, the question becomes how well-informed firms are about their industry price changes versus the fundamental. The following proposition shows that if strategic complementarity is large enough, then firms know more about the prices of their competitors than the fundamental, which holds only because they pay attention to their competitors mistakes.

Proposition 3. *In the equilibrium, if strategic complementarity is high enough, a firm knows more about the average prices of its competitors than about the fundamental and the aggregate price. A sufficient condition for this result is if* $\alpha\lambda \geq \frac{1}{2}$.[17]

This result is built solely on firms' incentives in paying attention to mistakes of their opponents. To see the reason, notice that the average price of a firm's competitors incorporates their average mistake:

$$p_{j,-k} = \delta q + v_{j,-k}.$$

[17] The necessary and sufficient condition in this sense has a complicated expression that is derived in the proof of the Proposition. It is shown that this result could hold even in occasions when $\alpha\lambda < \frac{1}{2}$ but K is small enough. For the purposes of this section, however, we only focus on this sufficient condition.

Hence, if a firm only paid attention to the fundamental, it would then know more about the fundamental than the prices of its competitors since their information would be orthogonal to the mistakes of others. It is only when the firm pays enough attention to $v_{j,-k}$ that it would know more about $p_{j,-k}$ than q.

This implies that when α is large, not only are prices more affected by the firms' average expectations of their own industries than by the fundamental, but these expectations are also formed under information sets that are more informative of these prices than they are of the fundamental. Therefore, poor expectations of the fundamental or aggregate prices are not necessarily an indicator of how well- or ill-informed firms are about their optimal prices or prices in their industry.

2.4 An AS-AD Framework and Non-Neutrality of Money

The closed form solution for the static model provides an intuitive framework for analyzing the real effects of a shock to the nominal demand. In this simple setup the aggregate demand curve is given by the fact that the deviations of aggregate price and output in the economy from their mean should add up to the shock to the aggregate demand, q:

$$p = -y + q.$$

Moreover, the equilibrium covariance of the aggregate price with q, as characterized in the previous section, implies the following aggregate supply curve.[18]

$$p = (2^{2\kappa} - 1)(1 - \alpha)y.$$

Figure (2) shows how these real effects work in a classic AS-AD graph. When a positive shock to q shifts the aggregate demand curve of the economy to the right, firms do not observe it perfectly. Instead, they observe a signal whose value is larger than its mean. From the perspective of any firm, however, such a realization for its signal can come from a combination of three independent sources: an increase in aggregate demand q, a common mistake of their industry in perceiving the realization of q, or an independent mistake on their own part in perceiving the value of q. The non-neutrality of money rises from firm's different incentives in responding to each of these possibilities. The degree to which a firm increases their price due to a change in q versus a common mistake in their industry is different, while they would rather not change their price at all in response to their own independent mistakes. These different incentives make firms reluctant in responding one to one to their realized signal:

[18] Aggregate supply can be derived from the equilibrium result $p = \delta q = \delta(p + y) \Rightarrow p = \frac{\delta}{1-\delta} y$.

they respond by a smaller magnitude due to the possibility that it may simply be a mistake, bearing in mind also the beliefs and responses of their competitors. As a result, when q goes up by one percent, firms across the economy increase their prices by less than that, on average. This creates an excess demand for goods, which increases the aggregate output.

The slope of the AS curve, which determines how an increase in q is divided between prices and output, depends on how much firms are capable of separating the three independent sources that affect their signals from one another. As κ increases, mistakes become smaller in the equilibrium as all firms see more informative signals of q. Therefore, if a firm sees a signal larger than its mean, they assign more probability to the case that the increase is coming from q rather than a mistake on their own or their competitors' part. Hence, they respond more strongly to their signals with a larger increase in their prices, which diminishes the effect of the shock on their output. Moreover, when strategic complementarity is smaller, the firms worry less about the mistakes of their competitors, and focus a higher amount of their attention on finding out the realization of q, which again diminishes the real effects of the shock. In the extreme case when $\kappa \to \infty$, signals are infinitely precise in revealing the realization of q and all firms respond one to one to their signals. This corresponds to an infinite slope for the AS curve where money is neutral, and output is completely unaffected by changes in the nominal aggregate demand.

3 Model Predictions and Relation to the Data

I use a unique quantitative survey of firms' expectations from New Zealand, which is comprehensively discussed in Coibion et al. (2015); Kumar et al. (2015). The survey was conducted among a random sample of firms in New Zealand with broad sectoral coverage. So far, six waves of the survey have been done over the time frame of September 2013 to July 2016.

Here, I only rely on the aspects of the survey that link the competitiveness of firms to the quality of their information. Motivated by the predictions of the model, I focus specifically on firms' quality of information about aggregates relative to their industries.

3.1 Number of Competitors and Strategic Complementarity

The two underlying assumptions of this paper that drive its main results are the finiteness of firms within industries and the existence of industry-level strategic complementarities. Two questions in the survey directly measure these for every firm within the sample and address these assumptions.

The first asks firms how many direct competitors they face for their main product or product line. The average firm in the sample reports that they face eight competitors, as documented in Table (1), with 35% of firms reporting that they four or fewer competitors. A breakdown of firms' answers from different industries shows that this average is fairly uniform across them. Column (3) of the table also reports a weighted average of firms' answers to this question based on their share of production in the whole sample. This weighted average aims to capture the representativeness of each firm in the economy, and the fact that it is is 5 shows that larger firms, with higher shares of production, face fewer competitors than the average firm in the sample.

To capture the degree of within industry strategic complementarity firms were asked the following question.

> "[S]uppose that you get news that the general level of prices went up by 10% in the economy:
>
> a. By what percentage do you think your competitors would raise their prices on average?
>
> b. By what percentage would your firm raise its price on average?
>
> c. By what percentage would your firm raise its price if your competitors did not change their price at all in response to this news?"

Given this hypothetical question, the pricing best response derived in the previous section allows us to back out the degree of strategic complementarity for each firm:

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

The average response of firms to this question is 0.9, which is relatively uniform across different industries. These responses indicate a high degree of strategic complementarity, which is in line with the standard calibrations of the analog of this parameter in the literature.[19]

3.2 Knowledge about Industry versus Aggregate Inflation

One of the main predictions of the model is that in the presence of coordination at the micro-level, firms are more aware of their industry 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. Figure (3) shows the distribution of firms' nowcasts of these two objects. While the average nowcast for

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

aggregate inflation, 4.3%, is very high and far from the actual inflation of 0.8%, the average nowcast of firms from their industry prices, 0.95%, is very close to this realized inflation. This observation directly parallels with the result in Corollary 1 that shows under imperfect competition prices are closer to average expectations of firms from their own industry prices than their average expectation of the aggregate price. Also, Table (2) reports the size of firms' nowcast errors in perceiving the two.[20] The average absolute nowcast error from industry inflation is 1.2%, a magnitude that is considerably lower than the average absolute nowcast error about aggregate inflation, 3.1%. This evidence is consistent with the prediction of Proposition 3 which states that in presence of high strategic complementarities firms are relatively more aware of their industry price changes than the aggregate ones. In addition, Figure (4) shows that on top of 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.

Firms' forecasts also exhibit similar patterns. Figure (5) shows the distribution of firms' average forecasts along with the forecast of the Reserve Bank of New Zealand (RBNZ) for the aggregate inflation from the sixth wave of the survey, conducted in second quarter of 2016. Despite a long history of low and stable inflation in New Zealand, a significant number of firms forecast a high rate of inflation for the following year, with an average of 3.6%. However, the distribution of forecasts for industry rate of inflation is both narrower and more symmetric around a mean forecast of 0.8%, which is very close to the range of inflation that New Zealand has been experiencing in the last few years.[21]

From the perspective of the standard models of inflation dynamics which relate the rate of inflation to firms' expectations of aggregate inflation, these high expectations of aggregate inflation seem very puzzling. Despite these large inflation forecasts among firms, which are consistently higher than the 2% target of the RBNZ in all waves of the survey, coupled with the fact that there were no significant changes in the output gap of New Zealand in this period, yearly inflation in New Zealand has been even lower than the target. Since the year 2012, yearly inflation has been averaging around 1%, with a high of 1.6% in the second quarter of 2014 and a low of 0.1% in the third quarter of 2015.

This evidence is also consistent with managers' subjective uncertainty of aggregate and industry-level inflation. In the sixth wave of the survey, firms were asked to assign proba-

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

[21] At the time of the survey, second quarter of 2016, yearly inflation was 0.8 percent, and RBNZ's forecast for the following year was 1.5%.

bilities to different outcomes regarding industry and aggregate yearly inflation.[22] Table (3) reports the standard deviation of managers' reported distribution for both of these objects, which I interpret as their subjective uncertainty. Firms are relatively less uncertain about their industry inflation in the following year than the aggregate one. This relates well to the prediction of the model in Proposition 3 that in the presence of high industry level strategic complementarity firms should know more about their industry than the aggregate economy as a direct implication of their strategic incentives in acquiring information.

3.3 Uncertainty about Inflation versus Number of Competitors

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.

$$\sigma_i^\pi = \beta_0 + \beta_1 K_i + \epsilon_i,$$

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 $\beta_1 < 0$. Panel (a) of Table (4) 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. The significance of this coefficient in explaining firms' uncertainty about aggregates is an observation that is not reconcilable neither with full information rational expectation models nor 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.

For comparison, I also run a similar regression of firms' uncertainty about their industry prices on their number of competitors:

$$\sigma_i^{\pi_i} = \beta_3 + \beta_4 K_i + \tilde{\epsilon}_i,$$

where now $\sigma_i^{\pi_i}$ is the standard deviation of firm i's reported distribution for their own industry. Panel (b) of Table (4) shows the result of and shows that they are also negatively correlated, yet with a smaller magnitude. This is also consistent with the model. As the number of a firm's competitors increase, firms become more certain about their price changes: in

[22] Firms were asked the following two 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." The bins to which firms assigned probabilities were identical in both questions.

larger industries mistakes wash out more effectively due to the law of large numbers, making the average price change more predictable. The smaller magnitude of the coefficient on the number of competitors, however, carries an important insight from the model. The same force that makes the prices of a firms' competitors more predictable, also discourages firms from paying attention to their mistakes. To see this from the model, recall that

$$p_{j,-k} = \delta q + v_{j,-k}.$$

We know that as the number of a firm's competitors goes up, their knowledge of the fundamental, q, also goes up, which is the result in Panel (a). However, this only happens because firms shift their attention from $v_{j,-k}$ to q, meaning that the decrease in uncertainty about the fundamental is accompanied by an increase in uncertainty about the mistakes of others. Hence, the decrease in uncertainty about the fundamental with the number of competitors should be lower in magnitude for industry prices than aggregate ones.

Panel (c) of Table (4) aims at capturing this effect by regressing the difference between firms' uncertainty about their industry relative to the aggregate inflation on the number of their competitors. This difference is positively correlated with the number of firms' competitors, consistent with the prediction that firms become relatively more uncertain about their industry price changes once the decline in uncertainty about the aggregates is extracted.

4 A Micro-founded Dynamic Model

This section extends the simple static model of Section 2 to a micro-founded dynamic general equilibrium model, whose ultimate purpose is to quantitatively analyze the effects of firms' strategic incentives in propagation of monetary policy shocks to aggregate output and inflation. All the derivations as well as the proofs for the propositions regarding the dynamic model are included in Appendix B.

4.1 Households

There is a large variety of goods produced in the economy. In particular, the economy consists of a large number of industries, $j \in J \equiv \{1, \ldots, J\}$; and each industry consists of $K \geq 2$ firms that produce weakly substitutable goods. The household takes the nominal prices of these goods as given and forms a demand over product of each firm in the economy.

In particular, the aggregate time t consumption of the household is

$$C_t \equiv \prod_{j\in J} C_{j,t}^{\frac{1}{J}}, \tag{3}$$

$$C_{j,t} \equiv \Phi(C_{j,1,t},\dots,C_{j,K,t}). \tag{4}$$

Where $C_{j,t}$ is the composite demand of household for the goods produced in industry j, and $\Phi(.) : \mathbb{R}^K \to \mathbb{R}$ is a continuously differentiable aggregation function that is homogeneous of degree 1, symmetric across its arguments, and such that $\Phi(1,\dots,1) = K$. Moreover, a specific form for $\Phi(.)$, which I will use to provide intuition in this section, is a CES aggregator with elasticity of substitution $\eta > 1$,

$$\bar{\Phi}(C_{j,1,t},\dots,C_{j,K,t}) = K\left(K^{-1}\sum_{k\in K} C_{j,k,t}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}.$$

Equation (3) denotes that the aggregate consumption of the household is Cobb-Douglas in the composite goods of industries. Also, household preferences over combining and consuming goods within industries, captured by the form of $\Phi(.)$, is central in determining the degree of within-industry strategic complementarity.[23]

Since 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.[24]

The representative household's problem is

$$\max_{((C_{j,k,t})_{(j,k)\in J\times K}, C_t, L_t, B_t)_{t=0}^{\infty}} \quad \mathbb{E}_0^f \sum_{t=0}^{\infty} \beta^t [\log(C_t) - \phi L_t] \tag{5}$$

$$s.t. \quad \sum_{(j,k)\in J\times K} P_{j,k,t}C_{j,k,t} + B_t \le W_t L_t + (1+i_{t-1})B_{t-1} + \sum_{(j,k)\in J\times K} \Pi_{j,k,t} - Tax, \ \forall t \ge 0$$

$$C_t = \prod_{j\in J} \Phi(C_{j,1,t},\dots,C_{j,K,t})^{J^{-1}}, \ \forall t \ge 0.$$

where $\mathbb{E}_t^f[.]$ is the full information rational expectations operator at time t, L_t is the labor supply of household, B_t is their demand for nominal bonds, W_t is the nominal wage, i_t is the

[23]The generality assumption on the form of function $\Phi(.)$ is mainly due to calibration purposes, as the CES aggregator is too restrictive in matching the level of strategic complementarity observed in the data. This generality assumption is not new to the literature of oligopolistic pricing in macroeconomic models. See, for instance, Rotemberg and Woodford (1992).

[24]While this might not be a very realistic assumption, it is the standard approach in the literature as a natural first step in separating the implications of rational inattention for households versus firms.

net nominal interest rate, $\Pi_{j,k,t}$ is the profits of firm j, k, and Tax is a constant lump sum tax that is used by the government to finance a hiring subsidy for firms that eliminates any long-run inefficiencies of imperfect competition.

I show in Appendix B that household's optimal behavior implies the following demand function for the product of firm j, k:

$$C_{j,k,t} = P_t C_t \mathcal{D}(P_{j,k,t}; P_{j,-k,t}) \tag{6}$$

where P_t is the price of the aggregate consumption bundle C_t, $P_{j,k,t}$ is firm j, k's price, and $P_{j,-k,t}$ is the vector of other firms' prices in sector j. Moreover, the function $\mathcal{D}(.;.)$ is homogeneous of degree -1, which in the case of CES aggregation reduces to

$$\bar{\mathcal{D}}(P_{j,k,t}, P_{j,-k,t}) = \frac{P_{j,k,t}^{-\eta}}{\sum_{l \in K} P_{j,l,t}^{1-\eta}}.$$

Let $Q_t \equiv P_t C_t$ be the aggregate nominal demand for the economy. Then, the household's intertemporal Euler and labor supply equations are given by:

$$\begin{aligned} 1 &= \beta(1+i_t)\mathbb{E}_t^f[\tfrac{Q_t}{Q_{t+1}}] \quad , \\ W_t &= \phi Q_t \qquad\qquad . \end{aligned}$$

The log-utility implies that the intertemporal Euler equation simply relates the level of nominal interest rate to the expected growth of the aggregate demand. This creates a natural duality between formulating monetary policy either in terms of the nominal interest rates, or specifying a law of motion for the aggregate demand, which is a well-known and frequently used result in the literature. Moreover, the linear disutility in labor, which corresponds to an infinite Frisch elasticity of labor supply, implies that the nominal wages are proportional to the nominal demand.[25]

4.2 Firms

Firms take wages and their demand from the household side as given and at each period set their prices based on their chosen information set up to that time; while committing to produce the realized level of demand that their price induces. Since my main objective is to examine the real effects of monetary policy through endogenous information acquisition

[25] The 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* industry strategic complementarity from the household side. I use this assumption to the same end in order to mainly focus on micro-founding *within* industry strategic complementarities.

of these firms, I abstract from other sources of monetary non-neutrality, and in particular assume that prices are perfectly flexible. After setting their prices, firms then hire labor from their own industry labor market and produce with a production function that is linear in their labor demand; $Y_{j,k,t} = L_{j,k,t}$. To eliminate the steady state inefficiencies of imperfect competition, I assume that there is a constant subsidy in the economy for hiring a unit of labor. Thus, firm j, k's nominal profit at time t is given by

$$\Pi_{j,k,t} = Q_t \Pi(P_{j,k,t}, P_{j,-k,t}, W_t),$$

where

$$\Pi(P_{j,k,t}, P_{j,-k,t}, W_t) \equiv (P_{j,k,t} - (1-\bar{\mathtt{s}})W_t)\mathcal{D}(P_{j,k,t}, P_{j,-k,t}).$$

Here, $P_{j,k,t}$ is firm j, k's own price, and $P_{j,-k,t}$ is the vector of other firms' price in industry j, Q_t is the nominal aggregate demand, W_t is the nominal wage, and $\bar{\mathtt{s}}$ is the hiring subsidy per unit of labor. We assume that there is a large number of industries in the economy so that every firm's effect on aggregate nominal demand is negligible.

Firms are *rationally inattentive*, and at each period t choose their information set optimally from a set of *available signals*, $\mathcal{S}^t$. In Appendix B, I carefully define these concepts for the dynamic model. Here, I focus on characterizing the firms' problem taking these definitions as given.

A strategy for any firm is to choose a set of signals to observe over time $(S_{j,k,t} \subset \mathcal{S}^t)_{t=0}^{\infty}$ and a pricing strategy that maps its information set to their optimal price at any given period, $P_{j,k,t} : S_{j,k}^t \to \mathbb{R}$, where $S_{j,k}^t = (S_{j,k,\tau})_{\tau=0}^t$ is the firm's information set at time t. Accordingly, given a strategy for all other firms in the economy, $(S_{j,l}^t \subset \mathcal{S}^t, P_{j,l,t} : S_{j,l}^t \to \mathbb{R})_{t=0, j \in J, l \neq k}^{\infty}$, firm j, k's problem is to maximize the net present value of their life time profits given an initial information set that they inherit at the time of maximization:

$$\begin{aligned} &\max_{(S_{j,k,t} \subset \mathcal{S}^t, P_{j,k,t}: S_{j,k}^t \to \mathbb{R})_{t=0}^{\infty}} \mathbb{E}[\sum_{t=0}^{\infty} \beta^t Q_0 \Pi(P_{j,k,t}(S_{j,k}^t), P_{j,-k,t}(S_{j,-k}^t), W_t) | S_{j,k}^{-1}] \qquad (7) \\ s.t. \quad & \mathcal{I}(S_{j,k,t}, (W_\tau, P_{l,m,\tau}(S_{l,m}^\tau))_{\tau=0,(l,m)\neq(j,k)}^t | S_{j,k}^{t-1}) \leq \kappa, \ \forall t \geq 0 \\ & S_{j,k}^{-1} \in \mathcal{S}^{-1} \text{ given.} \end{aligned}$$

where the constraint implies 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 by κ bits.

4.3 Monetary Policy and General Equilibrium

For simplicity, I assume that the monetary policy is set in terms of the growth of aggregate demand. Following the literature[26], I assume that this growth rate is an AR(1) process with a persistence of ρ:

$$\log(\frac{Q_t}{Q_{t-1}}) = \rho \log(\frac{Q_{t-1}}{Q_{t-2}}) + u_t. \tag{8}$$

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

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

a strategy profile for firms given an initial set of signals

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

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 (5)

2. Firms: given 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 Ω^F and Ω^H, $\{Q_t \equiv P_t C_t\}_{t=0}^{\infty}$ satisfies the monetary policy rule specified in Equation (8).

4. Markets clear:

$$\begin{aligned} &\text{Goods Markets:} && C_{j,k,t} = Y_{j,k,t} && ,\forall j,k \in J\times K, \\ &\text{Labor Markets:} && \textstyle\sum_{(j,k)\in J\times K} L_{j,k,t}^d = L_t^s && . \end{aligned}$$

4.4 Discussion of Incentives

The main challenge in characterizing the solution to the model lies in solving the firms' problems. To do so, I derive the second order approximation of firms' losses from sub-optimal pricing following the rational inattention literature, and assume that they minimize the expected net present value of these losses subject to the attention constraint.[27] At any

[26]See, for instance, Maćkowiak and Wiederholt (2009); Mankiw and Reis (2002); Woodford (2003a).

[27]Before the second order approximation, I show profit maximization is equivalent to minimizing these losses over time.

time, given a realization of $P_{j,-k,t}$ and Q_t, a firm's profit loss from charging a price $P_{j,k,t}$ is given by

$$\begin{aligned} L(P_{j,k,t}, P_{j,-k,t}, W_t) &\equiv \max_x \Pi(x, P_{j,-k,t}, W_t) - \Pi(P_{j,k,t}, P_{j,-k,t}, W_t) \\ &= (p_{j,k,t} - (1-\alpha)w_t - \alpha \frac{1}{K-1}\sum_{l\neq k} p_{j,l,t})^2. \end{aligned}$$

Here, small letters denote percentage deviations from the steady state, and α is the degree of industry level strategic complementarity which is now directly linked to the micro-foundations of the model.

4.4.1 The Source of Strategic Complementarity

Contrary to models of monopolistic competition where constant elasticity of demand implies a constant markup for firms over their marginal cost, an oligopolistic environment makes these markups codependent. When a firm in an industry changes their price, in essence, it is influencing the distribution of the demand across all firms in their industry. In other words, the elasticity of demand for firms within an industry 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 codependence:

$$P^*_{jk,t} = \underbrace{\mu(P^*_{j,k,t}, P_{j,-k,t})}_{\text{optimal markup}} \underbrace{\phi(1-\bar{\mathrm{s}})Q_t}_{\text{wage}},$$

where the optimal markup has the familiar expression in terms of the elasticity of a firm's demand, $\mu(P^*_{j,k,t}, P_{j,-k,t}) \equiv \frac{\varepsilon_D(P^*_{j,k,t}, P_{j-k,t})}{\varepsilon_D(P^*_{j,k,t}, P_{j-k,t})-1}$, and $\varepsilon_D(P_{j,k,t}, P_{j,-k,t}) \equiv -\frac{\partial Y_{j,k,t}}{\partial P_{j,k,t}}\frac{P_{j,k,t}}{Y_{j,k,t}}$ is firm j,k's elasticity of demand with respect to its own price.[28] In particular, in the case of the CES aggregator for industry goods this elasticity is

$$\varepsilon_D(P_{j,k,t}, P_{j,-k,t}) = \eta - (\eta-1)\frac{P_{j,k,t}^{1-\eta}}{\sum_{l\in K} P_{j,l,t}^{1-\eta}}. \tag{9}$$

This expression clarifies the source of industry level strategic complementarity in this economy. Firms lose profits if they do not adjust their markup due to changes in their competitors' prices. The expressions for strategic complementarity and markups in the

[28] The assumption that the aggregator function over industry goods, $\Phi(.)$, is homogeneous of degree one implies that demand elasticities and markups are independent of the level of nominal prices and solely depend on the relative prices of firms within an industry. In other words, $\mu(.,.)$ and $\varepsilon_D(.,.)$ are homogeneous of degree zero.

steady state are:

$$\alpha = \frac{1-\eta^{-1}}{K},\ \mu = \frac{\eta}{\eta-1} + \frac{1}{(\eta-1)(K-1)}.$$

The specific example of the CES aggregator shows that more competition in terms of the number of firms not only decreases the average markups of firms, which is a very intuitive implication of competition, but also decreases the strategic complementarity within industries. The reason for the latter is simple: as the number of competitors grows within an industry, every firm becomes smaller in proportion to its competitors and equally incapable of affecting their elasticity of demand by their price. In other words, when $K \to \infty$ we are in a state of monopolistic competition within industries, where the optimal markup of every firm only depends on how differentiated their good is from those of its competitors.

These expressions for the CES aggregator also show why this particular aggregator is too restrictive for a quantitative analysis of this model. Strategic complementarity under this aggregator is bounded above by 0.5 because $K \geq 2$, which is no way near the value of 0.9 that is observed in the data. To match this level of strategic complementarity quantitatively while simultaneously keeping the qualitative properties of this aggregator, I consider the following generalization. Suppose the household's preferences are such that their elasticities of demand for goods within industries are given by

$$\varepsilon_D(P_{j,k,t}; P_{j,-k,t}) = \eta - (\eta-1)\left(\frac{P_{j,k,t}^{1-\eta}}{\sum_{k\in K} P_{j,k,t}^{1-\eta}}\right)^{1+\xi} \underbrace{\left(\frac{\bar{P}^{1-\eta}}{\sum_{k\in K}\bar{P}^{1-\eta}}\right)^{-\xi}}_{=\frac{1}{K}},$$

where the new parameter ξ now captures how the elasticity of demand changes with the relative prices within the industry and allows us to match the elasticity of the markup independently.[29] Notice that this specification preserves the steady state properties of the CES aggregator up to the elasticities of demand and average markups, and only changes the elasticity of the elasticity of demand, which is related to the third order derivatives of the function $\Phi(.)$. It embeds the CES aggregator when $\xi = 0$, and the two are the same function

[29]Appendix B derives the demand functions that imply such elasticites. While the closed form solution for $\Phi(.)$ is not easy to derive, what we care about are these elasticities, and not the closed form of $\Phi(.)$ per se. An alternative method to match a higher degree of super-elasticity than that of the CES aggregator is to employ a more general aggregator as in Kimball (1995), a recent survey of whose applications is discussed in Gopinath and Itskhoki (2011) and Klenow and Willis (2016). I derive the demand function of the firms given a general form of such an aggregator in Appendix B, and discuss its implications for the degree of strategic complementarity. I show that while the Kimball aggregator allows for calibrating α to the level that is observed in the data, such a calibration leads to a counterintuitive result where the degree of strategic complementarity is decreasing with the elasticity of substitution across industry goods, meaning that a firm's profit depends less on the prices of its competitors when their goods become more substitutable. I depart from this aggregator by directly specifying the elasticity functions.

for all values of ξ when $K \to \infty$, which corresponds to to having a measure of firms within industries.[30]

Proposition 4. *There is strict industry level strategic complementarity in pricing, meaning that* $\alpha \in (0,1)$*, as long as a firm's elasticity of demand is increasing in their price, which corresponds to* $\xi > -1$*. Moreover, strategic complementarity is increasing in the elasticity of substitution* η*, decreasing in the number of firms within industries,* K*, and converges to zero as* $K \to \infty$*. The expression for* α *is*

$$\alpha = \frac{(1+\xi)(1-\eta^{-1})}{K + \xi(1-\eta^{-1})}.$$

The expression for α manifests how strategic complementarity is affected by both competition and the substitutability of goods. As the number of competitors for a firm goes up and their market share becomes smaller, their elasticity of demand loses its sensitivity to the price of the firm relative to its competitors. In the limit when $K \to \infty$, the firm takes this elasticity as given and realizes that its relative price has no effect on the demand curve of the industry in which they operate due to its small size. In this limit, the firm knows that the only demand that they face is only due to the weak substitutability of their good with respect to their competitors, and understand that their optimal price is just a constant markup over their marginal cost, independent of what their competitors do.

The elasticity of substitution, however, has the opposite effect on strategic complementarity. The higher is η, the more substitutable the firm's product is with those of its competitors implying that the firm would lose more profits if they do not match the optimal relative price with their competitors. Accordingly, more substitutability translates into higher strategic complementarity.

Finally, a larger value for the parameter ξ, which captures the sensitivity of the elasticity of a firm's demand to the relative prices in its industry, imply a larger degree of strategic complementarity. This parameter now allows us to match the super-elasticity of a firm's demand independently from the its demand elasticity.

4.4.2 Incentives in Information Acquisition

Appendix B thoroughly discusses my approach for solving the rational inattention problem of the firms. Here I discuss the dynamic and strategic incentives of firms in acquiring information.

[30] Notice that these elasticities are also well-defined in the sense that $\varepsilon_D(P_{j,k,t}, P_{j,-k,t}) \geq 1$ in a neighborhood around any symmetric point.

The specification of a firm's problem shows how their information set becomes the source of a dynamic trade-off. At every period, firms understand that the information that they choose to see will not only inform them about their contemporaneous optimal price, but also about their future payoffs. The following re-arrangement of a firm's problem manifests this trade-off:

$$\begin{aligned} V_t(S_{j,k}^{t-1}) \quad = \quad & \max_{S_{j,k,t}\subset\mathcal{S}^t, P_{j,k,t}(S_{j,k}^t)} \mathbb{E}[\underbrace{Q_0\Pi(P_{j,k,t}(S_{j,k}^t), P_{j,-k,t}(S_{j,-k}^t), \phi Q_t)}_{\text{contemporaneous payoff of } S_{j,k}^t} + \underbrace{\beta V_{t+1}(S_{j,k}^t)}_{\text{continuation value of } S_{j,k}^t} \quad |S_{j,k}^{t-1}] \\ s.t. \quad & \mathcal{I}(S_{j,k,t}, (Q_\tau, P_{l,m,\tau}(S_{l,m}^\tau))_{\tau=0,l,m\neq j,k}^t | S_{j,k}^{t-1}) \leq \kappa, \\ & \underbrace{S_{j,k}^t = S_{j,k}^{t-1} \cup \{S_{j,k,t}\}}_{\text{evolution of the information set}} \qquad . \end{aligned}$$

The dynamic incentives of a rationally inattentive agent is the main focus of Afrouzi and Yang (2016).[31] At every point in time the agent is choosing an information structure that serves two purposes: first, it gives them information about their optimal price at that time, which is the object that they want to learn about and match, and second, it forms a prior over their optimal prices in future periods. This implies that forward looking agents smooth out their mistakes over time by creating excess serial correlation in their information set. Intuitively, the manager of a firm wants to be informed of all their current and future optimal prices at a given point in time to the degree that it is feasible. Smoothing out mistakes means that the manager prefers an information set that compensates their current mistakes by ones in the future that move in the opposite direction. This way, for instance even if they make a mistake at a given time and charge a sub-optimally high price, their information set would help them to compensate for that by charging lower prices in the future.

The core difference between this paper and Afrouzi and Yang (2016) is the inclusion of the strategic trade-off that imperfectly competitive firms face in allocating their attention. When firms face competitors who make mistakes, they find it optimal to pay attention to those mistakes in addition to learning the fundamental. Their limited attention defines a possibility frontier for this attention allocation: an agent who wants to learn more about the fundamental is bound to learn less about the mistakes of its opponents and vice versa. In this sense, the objective in this paper is to find the optimal allocation of attention that takes place in the equilibrium. In particular, in an equilibrium where agents do pay attention to mistakes of one another, their mistakes become correlated. When the competitors of a firm in an industry charge a sub-optimally low price by mistake, the firm finds it optimal to pay

[31] We show in Afrouzi and Yang (2016) when such a problem is indeed a contraction mapping so that a unique $V(.)$ exists. Here we take that result as given.

attention to those mistakes and incorporate them in its price.

To better understand the separate roles that these strategic incentives of a firm plays in their price-setting behavior from its dynamic incentives, in this paper I shut down the dynamic incentives of firms completely by assuming that in their problem $\beta = 0$. The following Proposition derives the form of the optimal signals.

Proposition 5. *Given a strategy profile for all other firms in the economy, a particular firm prefers to see only one signal at any given time. Moreover, the optimal signal of firm* j,k *at time* t *is*

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

where q_t *is the nominal aggregate demand,* $p_{j,-k,t}$ *is the average price of* j,k*'s competitors, and* $e_{j,k,t}$ *is the rational inattention error of the firm.*

The closed form for the optimal signal in this case shows how firms incorporate the mistakes of their competitors into their information sets. To see this given the strategy of others, decompose their average price at time t to its projection on the history of fundamentals and the part that is orthogonal to them

$$p_{j,-k,t}(S^t_{j,-k}) = \underbrace{p_{j,-k,t}(S_{j,-k,t})|_q}_{\text{projection on realizations of all } q_{t-\tau}\text{'s}} + \underbrace{v_{j,-k,t}}_{\text{orthgonal to all realizations of } q_{t-\tau}\text{'s}}.$$

This is analogous to the decomposition that I did in the static model. It separates the average prices of others to a part that is linearly projected on current and past realizations of the fundamental, and a part that is orthogonal to it, denoted by $v_{j,-k,t}$. Similar to before, we call these the mistakes of a firm's competitors in pricing. Notice that the finiteness of the number of competitors immediately implies that $var(v_{j,-k,t}) \neq 0$. Given this decomposition, the optimal signal of the firm is

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

This decomposition of the signal illustrates the main departure of this paper from models that assume a measure of firms. Since $var(v_{j,-k,t}) \neq 0$, the signal of a firm co-varies more with the price changes of its competitors than with the fundamentals of the economy.[32] When

[32] This in itself does not mean that the signal is more predictive of a firm's competitors' prices than the aggregate economy since predictive power of a signal also depends on the volatility of the variable that is being predicted, and industry prices are more volatile than the aggregate economy. However, as we showed

there is a measure of firms, however, the term $\alpha v_{j,-k,t}$ disappears and these two covariances converge to one another. Intuitively, going back to the result in Proposition 3, this implies that when α is large enough, and there are a finite number of firms in industries, firms are more informed about their own industry prices than the fundamentals of the economy. In the next subsection, I show how for large α's, it is these expectations that mainly drive the inflation in the economy.

Moreover, given the joint stochastic process of these signals, the best pricing response of a firm reduces to a Kalman filtering problem, which then implies that

$$p_{j,k,t}(S^t_{j,k,t}) = \sum_{\tau=0}^{\infty} \delta_\tau S_{j,k,t-\tau},$$

where $(\delta_\tau)_{\tau=0}^{\infty}$ is a summable sequence.[33] It is worth mentioning that these pricing strategies are not necessarily time independent, as the initial signal structure of firms determines their initial prior about the state of the economy, and affects their prices for periods to come. To get around this issue, in solving the model, I assume that the initial signal structure is such that these firms' best pricing responses are stationary. This is equivalent to assuming that the game starts with an information structure that corresponds to the steady state of firms' attention allocation problem. Intuitively, these δ_τ's represent the confidence of the agent on how informative each element of her information set is about the optimal price that she would like to charge at time t. If a firms did not make mistakes, then the only signal that would matter for it at time t would be $S_{j,k,t} = (1-\alpha)q_t + \alpha p_{j,-k,t}(S^t_{j,-k})$, so that $\delta_0 = 1, \delta_\tau = 0, \forall \tau \geq 1$. However, making mistakes over time reduces the informativeness of a firm's signals and it finds it optimal to put some weight on their previous signals in setting their prices. Therefore, the more uninformative the signals, the more persistent the response of firm's prices would be to a shock over time.

Given the result in this Proposition 5, solving the model reduces to finding the following fixed point: a symmetric stationary equilibrium is a stationary joint stochastic process for signals of firms, and a pricing strategy $(\delta^*_\tau)_{\tau=0}^{\infty}$, such that for any firm whose competitors set their prices according to this sequence, the firm finds it optimal to use $(\delta^*_\tau)_{\tau=0}^{\infty}$ for setting its prices. I solve this problem computationally by truncating the pricing strategy such that $|\delta^*_\tau| = 0, \forall \tau > T$, where T is large enough to verify that the stationary effects of the shock disappear before T periods, up to a computational tolerance.[34] Given the truncated pricing

in Proposition 3, once α gets large enough, this difference is large enough so that firms end up with more information about their own industry price changes than the fundamental.

[33] In general, these pricing strategies are time dependent; however, I will focus on an initial signal structure that supports a stationary equilibrium. For details, see Appendix B.

[34] I justify this method by showing that in equilibrium, $(\delta^*_\tau)_{\tau=0}^{\infty}$ is summable. Intuitively, firms with

strategy, I find the implied stationary joint stochastic process of signals for firms, characterize the optimal pricing strategy given this distribution, and iterate until convergence.

4.4.3 Inflation Dynamics and the Phillips Curve

The following Proposition derives the Phillips curve of this economy.

Proposition 6. *The Phillips curve of this economy is*

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

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

The Phillips curve illustrates the main insight of this paper. In economies with high industry level strategic complementarity (α close to 1), it is firms' average expectation of their own-industry price changes that drives aggregate inflation rather than their expectations of the growth in aggregate demand.[35] Moreover, Proposition 5 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; an insight that is comparable with Proposition 3 in the static model. Therefore, when α is large, not only is inflation driven more by firms' expectations of their own industry price changes, but also firms' expectations are formed under information structures that are more informative about their own industry price changes.

Additionally, the slope of the Phillips curve shows how these strategic complementarities, as well as the capacity of processing information, affect monetary non-neutrality in this economy. Higher capacity of processing information makes the Phillips curve steeper, such that in the limit when $\kappa \to \infty$, the Phillips curve is vertical. When firms have infinite attention, their estimates of the fundamental as well as their competitors prices are also infinitely precise. Firms immediately realize changes in the fundamental and react to it

positive capacity of processing information do not use signals that only inform them about a distant past of the economy.

[35]For comparison, we show in Afrouzi and Yang (2016) that in an economy with a measure of firms the Phillips curve is in terms of the firms' expectations of the aggregate inflation: $\pi_t = (1-\tilde{\alpha})\overline{\mathbb{E}_{t-1}[\Delta q_t]} + \tilde{\alpha}\overline{\mathbb{E}_{t-1}[\pi_t]} + (1-\tilde{\alpha})(2^{2\kappa}-1)y_t$, where $\tilde{\alpha}$ is the degree of across industry strategic complementarity. This is also comparable to a Phillips curve with sticky information: $\pi_t = (1-\tilde{\alpha})\overline{\mathbb{E}_{t-1}[\Delta q_t]} + \tilde{\alpha}\overline{\mathbb{E}_{t-1}[\pi_t]} + (1-\tilde{\alpha})\frac{\tilde{\lambda}}{1-\tilde{\lambda}}y_t$, where $\tilde{\lambda}$ is the fraction of firms that update their information in a given period. Notice that contrary to the result in this paper, in both these Phillips curve inflation is directly related to firms' average expectations of aggregate inflation.

under the common knowledge that every other firm is also doing so, which leads to complete monetary neutrality in the economy.

4.5 Calibration

Recall that the rational inattention problem of an industry is characterized by the following parameters: capacity of processing information for every firm, $\lambda = 1 - 2^{-2\kappa} \in [0, 1)$; the number of firms in the industry, K; and the degree of strategic complementarity α. Among these, the first two are deep parameters of the model; but α is pinned down as a function of ξ, η and K by the expression in Proposition 4.

Table (5) shows the calibrated values of these parameters. I calibrate the elasticity of substitution within industry goods, η, to a value of 6 to match the average markup of 30%, reported by the firms in the survey. A value of 6 for this parameter is also in line with the usual calibration in the macroeconomics literature. Moreover, I set the number of competitors in industries to a baseline value of 5, the average value reported by the firms weighted by their market share in the sample. Finally, I calibrate the curvature of the elasticity of demand, ξ, to 40 in order to match an average strategic complementarity of 0.9 as observed in the data. In addition, I calibrate the persistence of the growth in nominal demand, ρ, to the persistence of the nominal GDP growth in New Zealand, 0.5.[36]

Calibrating the capacity of processing information has been a challenge in the rational inattention literature due to a lack of suitable data so far. However, the New Zealand survey allows me to calibrate this parameter by directly measuring the quality of firms' information about aggregate inflation. To do so, I exploit the fact that λ is Kalman gain of firms in predicting their optimal prices, meaning that it is the weight that they put on their new information in predicting their optimal prices. I follow Coibion and Gorodnichenko (2015) to measure the degree of information rigidity in forecasts of aggregate inflation from the data, and then taking the calibration of other parameters as given I find $\lambda = 0.7$ generates the same degree of rigidity in firms' forecasts of aggregate inflation in the model.

A value of 0.7 is relatively large and represents a small degree of information rigidity, especially compared to the current models of noisy information, which usually assume calibrations that imply lower Kalman gains. The empirical literature has also estimated values that are less than 0.7. For instance, Coibion and Gorodnichenko (2015) estimate a Kalman gain of 0.5. This model, however, does not need a low value for λ to match the high degree of aggregate information rigidity due to its endogenous propagation mechanism. Despite a high λ at the micro level, firms spend a large portion of their attention tracking the mistakes

[36] 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.

of their competitors and the portion that is allocated to tracking aggregate fundamentals is therefore significantly lower. Simply put, firms devote a lot of attention to tracking their optimal prices, even more than what professional forecasters do for inflation. However, since they do not directly care about aggregate inflation, their forecasts manifests a high degree of rigidity.

5 Simulations and Counterfactuals

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 the propagation of monetary policy shocks to inflation and output through these incentives. Figure (6) shows the impulse responses of inflation and output to a one percent shock to the growth of nominal demand. The intuition behind non-neutrality of money is similar to the static case with the major difference being the persistence of real effects in the dynamic model. When the shock hits the economy, firms do not directly observe the shock itself, but instead see a signal. From the perspective of the firms, however, a large realization for their signal could also be due to a rational inattention error, as a result of which they react to the news reluctantly and increase their prices by less than the increase in nominal demand, leading to higher production. However, in contrast to the static model, which is equivalent to the dynamic model when q_t is i.i.d. over time, the effect of the shock persists over time because firms know that a shock to the fundamental continues affecting their optimal prices in later periods. Overtime, as firms keep observing more signals, they become more certain of both the source of the increase in their signals and the magnitude of its effect on their optimal price. After enough time passes, firms adjust their prices completely so that the real effects of the shock disappears.

The figure also shows how higher levels of competition affect the non-neutrality of money in a very significant fashion, such that doubling competition from its baseline calibration decreases the half-life of output response to the shock by 40% and reduces its on impact response by 15%. The effects on the response of inflation are similarly profound. A two-fold increase in the number of competitors at the micro-level reduces the half-life of inflation response by 25% and decreases its on impact response by 33%. These large quantitative effects of competition reflect two distinct channels through which competition alters economic outcomes. The first channel is the *attention reallocation* effect that competition has in the optimal attention allocation of firms given a fixed degree of strategic complementarity. As the number of competitors increase and the law of large numbers start to hold for the mistakes of a firm's competitors, the firm worries less about those mistakes and shifts their attention to tracking the fundamental shocks. Moreover, the dependence of α to K implies that the number of firms in industries also changes the equilibrium distribution of prices through a second channel, to which I will refer as the *strategic complementarity* effect. Higher competition alters the super-elasticities of firms' demand function and eliminates the dependence of their profits to the prices of other firms.

While these two channels affect the impulse response functions of the model in the same direction, from an economic perspective they are different in nature. The reallocation channel is novel to the literature and characterizes an effect that has been absent in previous models due to the assumption that every firm interacts with a measure of others, which in this paper correspond to the case where $K \to \infty$. The strategic complementarity channel is also new in the sense that it micro-founds the dependence of strategic complementarity to the number of competitors within every industry, but the effects of different levels of strategic complementarity on the propagation of monetary policy shocks in models of information rigidity has already been pointed out in the literature by seminal work of Mankiw and Reis (2002); Woodford (2003a); Maćkowiak and Wiederholt (2009). Therefore, the contribution of this paper in pointing out this later channel is mainly linking the strategic complementarity to the number of firms by micro-founding it. For the rest of this section I focus on both these channels and analyze the impulse responses of the model for each of them separately.

5.1 The Reallocation Effects of Competition

For several values of K, Figure (7) shows the impulse responses of inflation and output in the economy to a one percent shock to the growth of aggregate demand, fixing the value of strategic complementarity to its baseline calibration value of 0.9. By fixing the value of strategic complementarity, here I have shut down the strategic complementarity channel and the impulse responses represent only the reallocation effects of competition. Again, higher number of firms within industries corresponds to a less persistent output response in the economy, such that doubling competition from its baseline value decreases the half life of output by 18%, a little less than half of the 40% reduction in half-life under overall effects.

The reason for this relates to equilibrium incentives of firm in allocating their attention within the industries. When the economy is less competitive – K is small – firms are more worried about the mistakes of their competitors and allocate a high amount of attention to tracking those mistakes. Since mistakes are orthogonal to the elements of the fundamental, when all of the firms in the economy spend more resources to learning the mistakes of their competitors, they know less about the fundamental. The incentive of learning others' mistakes diminishes with the size of industries. When every firm in the economy competes with a large number of competitors, it is confident that the mistakes of others wash out, and allocates more attention to learning the fundamental of the economy. As a result, in a more competitive economy, firms pay more attention to the fundamental and learn it more quickly than firms in a less competitive economy. This manifests itself in two dimensions: when firms pay less attention to the fundamental over time by paying more attention to the mistakes of their competitor, it takes them *more time* to learn the fundamental through their signals. Thus, it takes longer for such firms to learn the shock and perfectly adjust their prices with respect to it, which then directly leads to a higher persistent response of output to the shock.

It is also worth mentioning that the effects of competition through this channel are inherent to the dynamic model and are completely absent in the static models of coordination. Recall from Section 2 that the average response of output in the static model is independent of K and is equal to $\frac{\lambda-\alpha\lambda}{1-\alpha\lambda}$. This result carries on to the dynamic model in terms of a K-independent *on-impact* response of output and inflation to the shock. At first glance to the reader, this might seem contradictory with the claim that more competitive industries respond more confidently to monetary policy shocks and manifest lower monetary non-neutrality.

This calls for a more detailed discussion of the firms' incentives within industries. Notice that firms not only respond to fundamental shocks in the economy, but also to the endogenously driven mistakes of their rivals. At any point in time, when a firm receives a signal with a high value, three possibilities occur to them: this may simply be a mistake that has occurred to them and only them, or it may be a *common* mistake across their industry, or it might be the case that a shock to the fundamental has happened in the economy. The key is that it is only in the first case that a firm would be unwilling to respond to this signal, while they would optimally want to respond to either of the later cases. While less competitive industries receive continuously less informative signals of the fundamental, their signals are instead more informative of the mistakes of others. In other words, it is always true that a firm in a less competitive industry, upon receiving any of its signals, is less sure of the changes in the fundamental. Such a firm, however, is more sure of the fact that it might be a common mistake in their industry, to which they would also respond by increasing their price. As a result, in the context of an impulse response function, when a firm sees a high signal after a long time of seeing zero signals, the on impact response of the firm will be independent of how competitive their industry is: independent of whether it is due to a shock to fundamental or a common mistakes, their response is to increase their price.

It is only *over time* that a firm is able to differentiate between these two cases due to the fact that shocks to the fundamental are more persistent than the endogenously driven mistakes.[37] When a firm keeps getting high signals over time, they assign more probability to the case that the shock was to the fundamental because if it was a common mistake, it should have had disappeared more quickly from their signals. The following Proposition formally shows that on impact response of output and inflation to a shock is independent of the degree of the competitiveness within industries.

Proposition 7. *When $\beta = 0$, the on impact response of output and inflation in the economy is independent of the reallocation effects of imperfect competition and is always given by*

$$\pi_0 = \tfrac{1-\alpha}{1-\lambda\alpha}\lambda \quad ,$$
$$y_0 = \tfrac{1-\lambda}{1-\lambda\alpha} \quad .$$

This result also points toward a fundamental difference between this model and one that incorporates a measure of firms. It is a feature of rational inattention models that a lower capacity of processing information or a higher degree of strategic complementarity increases the persistence of output response.[38] However, in those models, such effects are always accompanied by a lower on impact response of output. In other words, from an estimation or calibration perspective, those models introduce a trade-off between matching the on impact response of output and matching the persistence of that response over time. The model presented here, however, introduces a new micro-founded degree of freedom for separately matching these two moments.

5.2 The Strategic Complementarity Channel

For several values of K, Figure (8) shows the impulse responses of inflation and output in the economy to a one percent shock to the growth of aggregate demand only in response to the strategic complementarity channel. I shut down the reallocation channel by assuming that there are infinite firms in every industry, but exogenously impose a level of strategic complementarity that is implied by different levels of K.[39] As pointed out in the previous literature, higher strategic complementarity significantly changes the non-neutrality of money, such that doubling competition from its baseline value reduces the half-life of output by 22%, a little more than half of the overall effects in persistence. Moreover, it accounts for all of the decline in on-impact response of output as implied by Proposition 7.

[37] A shock to the fundamental will change the level of a firm's price forever as the fundamental has a unit root. However, mistakes vanish over time because firms have positive capacity. Hence, mistakes are less persistent than fundamental shocks.

[38] See, for instance, Afrouzi and Yang (2016).

[39] The implied values of strategic complementarity for $K = 5, 10$ and $K \to \infty$ are $0.9, 0.8$ and 0, respectively.

To discuss the intuition behind this effect recall from Proposition 5 that the optimal signal of the firm is a combination of current fundamental and the average prices of its competitors. Although we have shut down the reallocation affects though assuming that a firm's competitors do not make mistakes, the optimal signals of firms still hold this form. Since firms are still rationally inattentive, their optimal pricing strategy puts some weight on previous signals. This means that when there is higher degrees of strategic complementarity, instead of tracking the current value of the fundamental, the firm's signal puts a lot of weight on previous shocks to the fundamental because their competitors prices are affected by those previous shocks, which makes their signal less informative of the contemporaneous shock. As a result, when the firms see a high-valued signal in an economy with a higher degree of strategic complementarity, they are less confident that the signal is reflecting a contemporaneous shock to the fundamental, and they respond more reluctantly and less aggressively than what they would under a lower degree of strategic complementarity, which explains the change in the on-impact response of output for different values of α.

The intuition behind the change in the persistence of the responses is similar. When signals continuously put less weight on contemporaneous shocks, it takes firms longer to become fully informed about a shock, and therefore longer for the effects of the shock to disappear.

6 Concluding Remarks

> *"[the] complication is that we do not know whose expectations 'matter' for determining inflation."* Janet Yellen(2015)

Managing aggregate inflation expectations has been at the center of monetary policy makers' attention not only for controlling inflation but also as a potential instrument after the onset of the zero lower bound during the Great Recession. However, the expectations of price setters from aggregate inflation are highly biased and volatile in countries that have had low and stable inflation for decades, which goes against the close relationship that baseline monetary models predict between the two. Not only do these unanchored inflation expectations pose a serious challenge in reconciling standard models with the empirical evidence, but also render the unconventional monetary policies that aim on managing them ineffective.

In this paper, I develop a model to address this puzzle and show that what matters mainly for price setters is their expectations of their own industry inflation rather than aggregate inflation. Managers of firms do not directly care about aggregate inflation and are mainly concerned with how their own competitors change their prices in the face of a shock. In fact, when allowed to choose their information structure, managers are willing to sacrifice information about the aggregate economy by shifting their attention towards learning their competitors' prices. As a result, they are more informed about their optimal prices than what their expectations of aggregate inflation would suggest.

Moreover, I show that these endogenous informational incentives have significant implications for the propagation of monetary policy shocks. A two-fold increase in the number of competitors that a firm faces at the micro level decreases the half-life of output and inflation responses to a monetary policy shock by 40 and 25 percents respectively. The on impact effects are similarly large. Doubling the number of competitors for every firm reduces the on-impact response of output and inflation by 15 and 33 percents respectively.

The results of this paper provide valuable insights for policy makers. On the one hand, the fact that aggregate inflation is not the primary concern of 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 in doing so, it eliminates it as a concern for economic agents. 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. On the other hand, this also implies that managing expectations of aggregate inflation is neither an objective nor necessarily a powerful instrument for monetary policy. These expectations are relatively unimportant for firms and do not have much impact on their pricing decisions.

This result does not necessarily rule out policies that target managing expectations, but rather provides a new perspective on how those policies should be framed. 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 but also convince them that their competitors are also on board with these changes. 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 them in terms of how their industries would be affected. How policy can achieve these ends remains a question that deserves more investigation. However, in light of the results in this paper, the future work needs to focus more on how the policy would affect firms' expectations of their own industry price changes rather than the aggregate ones.

References

Afrouzi, H. and C. Yang (2016). Inflation dynamics under dynamic inattention. Mimeo.

Angeletos, G.-M. and C. Lian (2016). Incomplete information in macroeconomics: Accommodating frictions in coordination. Technical report, National Bureau of Economic Research. Working Paper No. 22297.

Angeletos, G.-M. and A. Pavan (2007). Efficient use of information and social value of information. *Econometrica 75*(4), pp. 1103–1142.

Bryan, M. F., B. Meyer, and N. Parker (2015). The inflation expectations of firms: What do they look like, are they accurate, and do they matter? Technical report, Federal Reserve Bank of Atlanta. Working Paper No. 2014-27a.

Cavallo, A., G. Cruces, and R. Perez-Truglia (2014). Inflation expectations, learning and supermarket prices: Evidence from field experiments. Technical report, National Bureau of Economic Research. Working Paper No. 20576.

Coibion, O. and Y. Gorodnichenko (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy 120*(1), 116–159.

Coibion, O. and Y. Gorodnichenko (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *The American Economic Review 105*(8), 2644–2678.

Coibion, O., Y. Gorodnichenko, and S. Kumar (2015). How do firms form their expectations? new survey evidence. Technical report, National Bureau of Economic Research. Working Paper No. 21092.

Colombo, L., G. Femminis, and A. Pavan (2014). Information acquisition and welfare. *The Review of Economic Studies 81*(289), 1438–1483.

Cover, T. M. and J. A. Thomas (2012). *Elements of Information Theory*. John Wiley & Sons.

Denti, T. (2015). Games with unrestricted information acquisition. Technical report. Mimeo.

Friedman, M. (1968). The role of monetary policy. *The American Economic Review 58*(1), 1–17.

Golosov, M. and R. E. Lucas Jr (2007). Menu costs and phillips curves. *Journal of Political Economy 115*(2), 171–199.

Gopinath, G. and O. Itskhoki (2011). In search of real rigidities. In *NBER Macroeconomics Annual 2010, Volume 25*, pp. 261–309. University of Chicago Press.

Hellwig, C. and L. Veldkamp (2009). Knowing what others know: Coordination motives in information acquisition. *The Review of Economic Studies 76*(1), 223–251.

Kimball, M. S. (1995). The quantitative analytics of the basic neomonetarist model. *Journal of Money, Credit and Banking 27*(4), 1241–1277.

Klenow, P. J. and J. L. Willis (2016). Real rigidities and nominal price changes. *Economica 83*(331), 443–472.

Kumar, S., H. Afrouzi, O. Coibion, and Y. Gorodnichenko (2015). Inflation targeting does not anchor inflation expectations: Evidence from firms in new zealand. *Brookings Papers on Economic Activity Fall*, 151–208.

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

Mackowiak, B., F. Matejka, and M. Wiederholt (2016). The rational inattention filter. Mimeo.

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

Maćkowiak, B. and M. Wiederholt (2015). Business cycle dynamics under rational inattention. *The Review of Economic Studies 82*(4), 1502–1532.

Mankiw, N. G. and R. Reis (2002). Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics 117*(4), pp. 1295–1328.

Morris, S. and H. S. Shin (2002). Social value of public information. *The American Economic Review 92*(5), pp. 1521–1534.

Myatt, D. P. and C. Wallace (2012). Endogenous information acquisition in coordination games. *The Review of Economic Studies 79*(1), 340–374.

Phelps, E. S. (1967). Phillips curves, expectations of inflation and optimal unemployment over time. *Economica*, 254–281.

Rotemberg, J. J. and M. Woodford (1992). Oligopolistic pricing and the effects of aggregate demand on economic activity. *The Journal of Political Economy 100*(6).

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System technical Journal 27*, 379–423.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Woodford, M. (2003a). Imperfect common knowledge and the effects of monetary policy. *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, 25.

Woodford, M. (2003b). *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press.

7 Figures

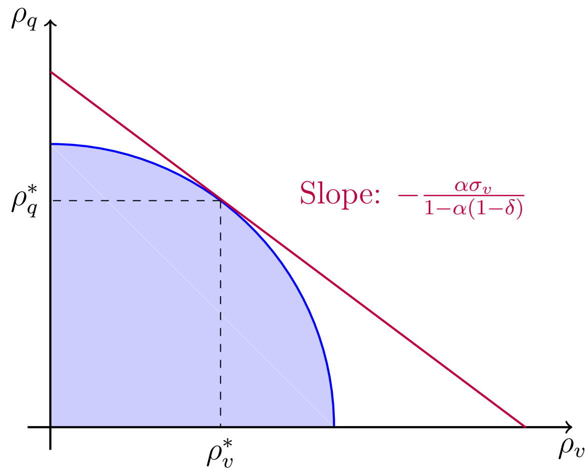

Figure 1: The figure symbolically illustrates the optimal correlations that a firm chooses among their signal, the fundamental q, and the mistakes of their competitors, v.

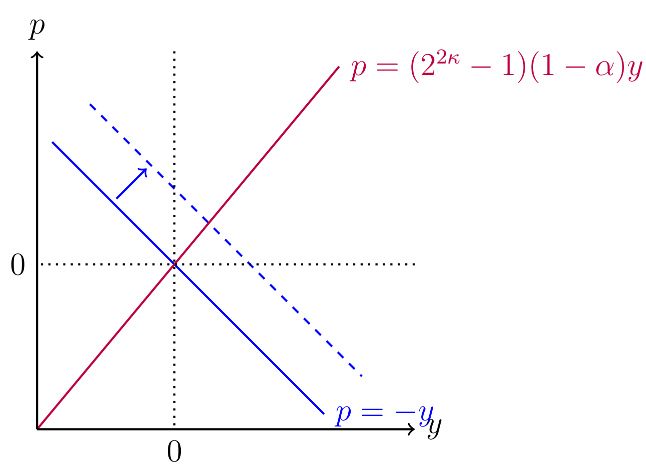

Figure 2: The figure depicts the real effects of a shock to a nominal demand within the static model. The magnitude of the real effect decreases with the capacity of processing information and increases with the degree of strategic complementarity.

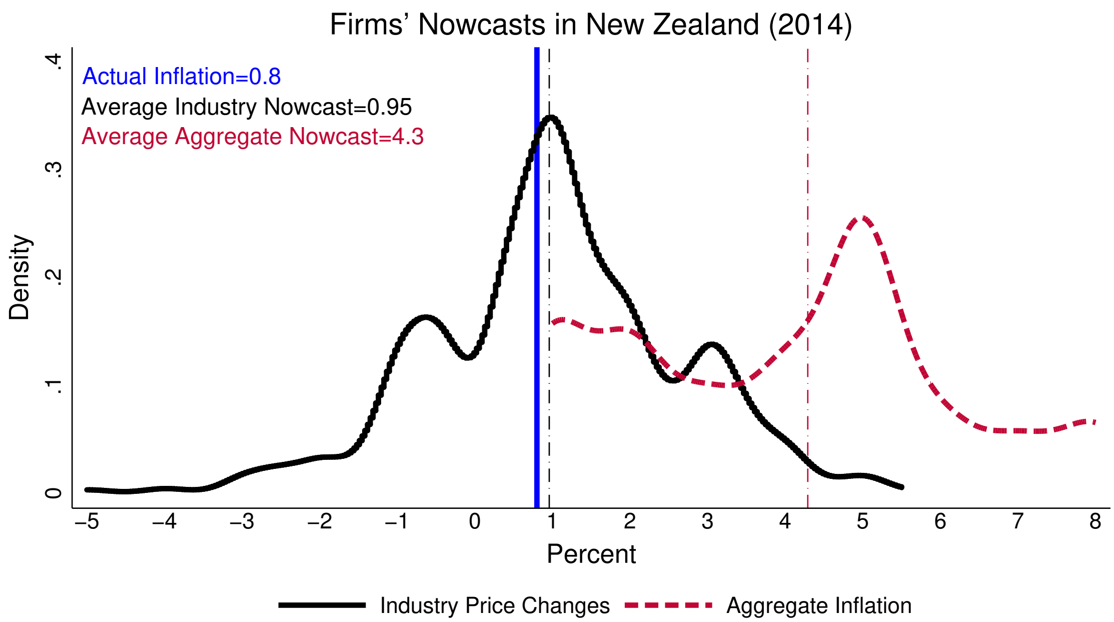

Figure 3: The figure shows the distributions of firms' nowcasts for both aggregate and industry level inflation. The dashed vertical lines show the means of these distributions. The solid vertical line shows the realized inflation that firms were nowcasting.

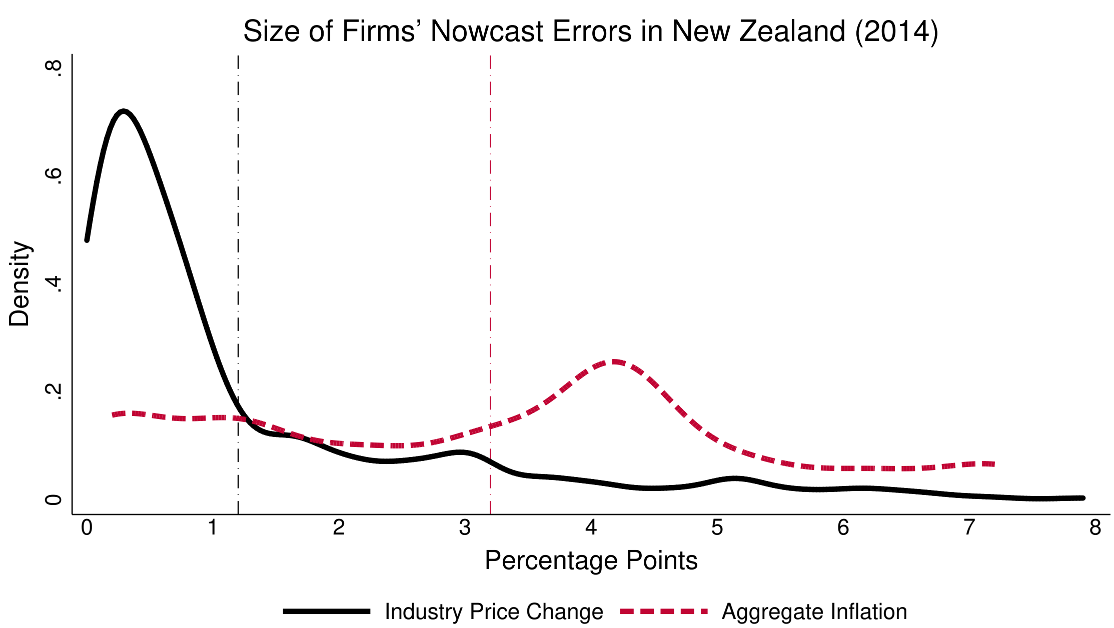

Figure 4: The figure shows the distributions of the size of firms' nowcast errors for aggregate and own-industry inflation. The dashed vertical lines show the means of these distributions.

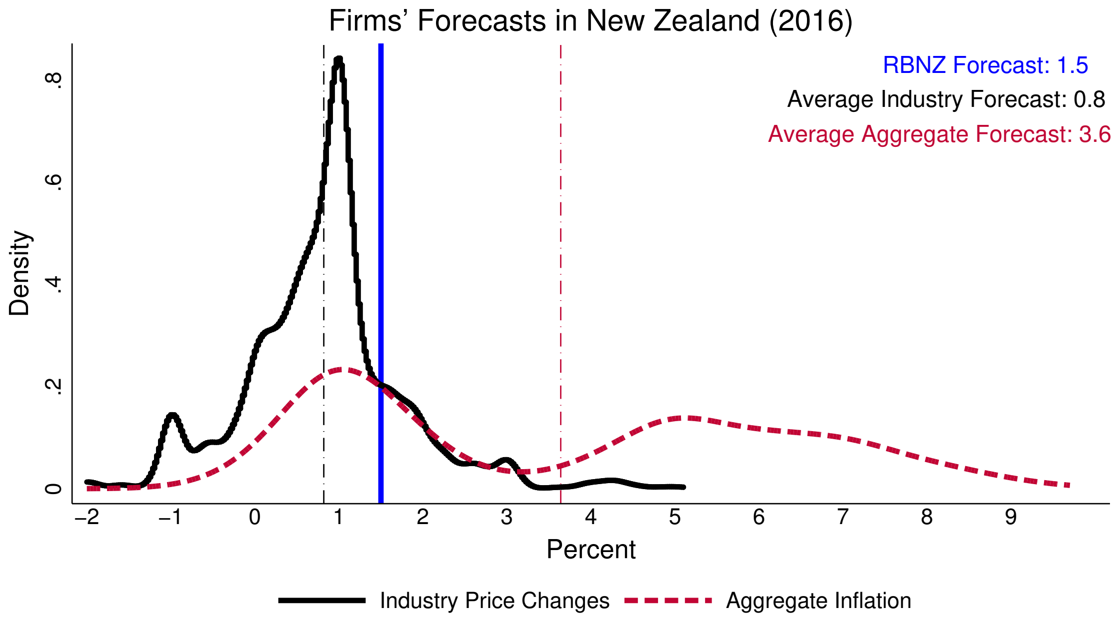

Figure 5: The figure shows the distribution of firms' forecast for aggregate and own-industry inflation. The dashed vertical lines denote the means of these distributions. The solid vertical line shows the RBNZ's forecast of aggregate inflation for the same horizon that firms were forecasting.

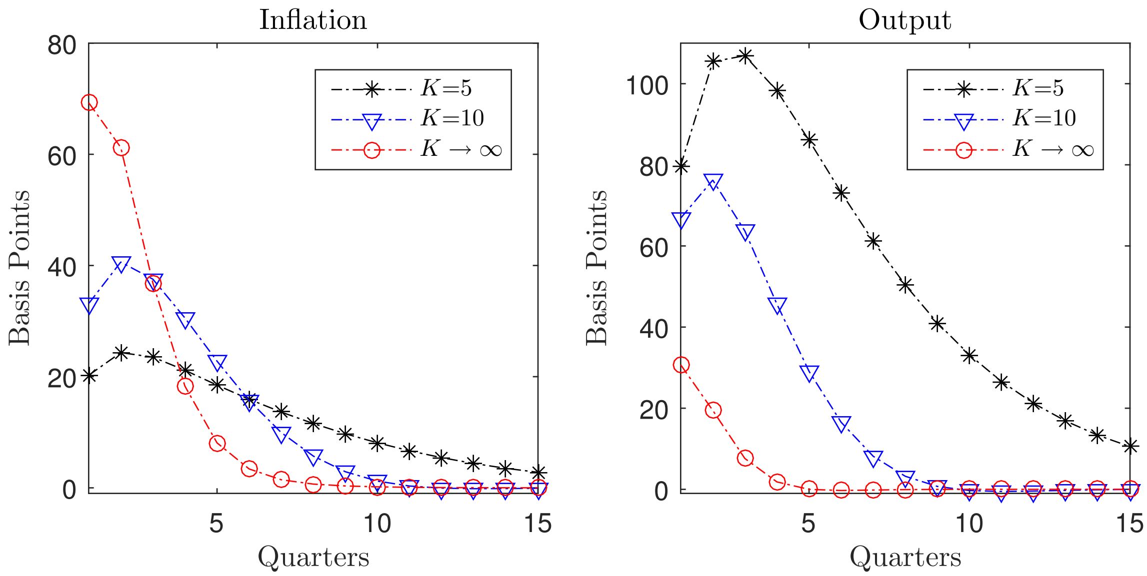

Figure 6: The figure shows impulse response functions of output and inflation to a 1 percent shock to the growth of aggregate demand, for overall effects of different values of K. More competitive economies respond more strongly and more quickly to the shock. See Section 5 for a discussion of these results.

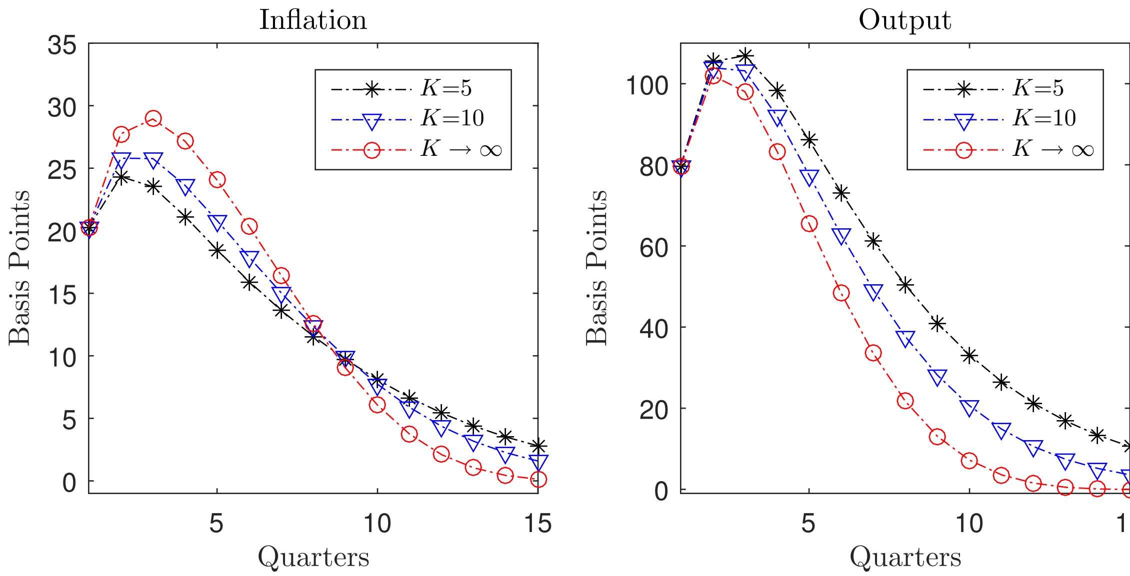

Figure 7: The figure shows impulse response functions of output and inflation to a 1 percent shock to the growth of aggregate demand, for *reallocation* effects of different values of K. More competitive economies respond more strongly and more quickly to the shock. See Section 5 for a discussion of these results.

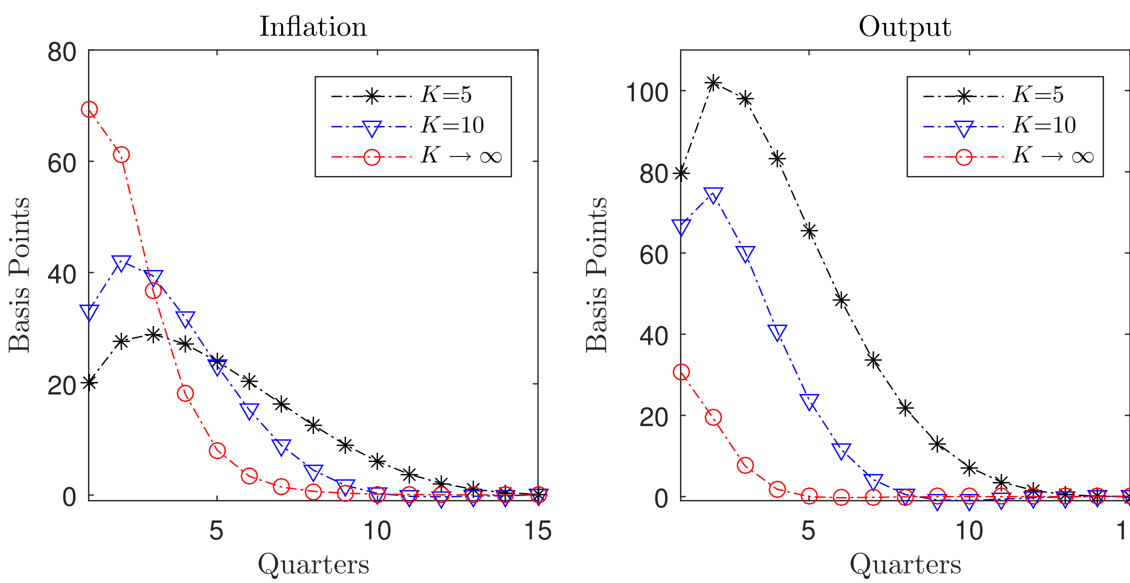

Figure 8: The figure shows impulse response functions of output and inflation to a 1 percent shock to the growth of aggregate demand, for *strategic complementarity* effects of different values of K. More competitive economies respond more strongly and more quickly to the shock. See Section 5 for a discussion of these results.

8 Tables

Table 1: Summary Statistics for Number of Competitors and Strategic Complementarity

		Number of competitors[a]		*Strategic complementarity*[b]	
	Observ- -ations	*Mean (Std. Dev.)*	*Weighted mean*	*Mean (Std. Dev.)*	*Weighted mean*
Industry	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Construction and Transportation	289	6.8	4.6	0.93	0.93
		(4.9)		(0.34)	
Manufacturing	715	8.2	4.9	0.87	0.96
		(6.4)		(0.37)	
Professional Financial Services	617	8.6	5.0	0.93	0.92
		(6.5)		(0.28)	
Trade	419	9.0	4.7	0.90	0.91
		(6.3)		(0.32)	
Total	2040	8.2	4.8	0.91	0.94
		(6.0)		(0.33)	

a. Column (2) reports the average number of competitors along with standard deviations. Column (3) reports the average number of competitors weighted by firms' share of total production in the sample.
b. Column (4) reports the average strategic complementarity along with standard deviations. Column (4) reports the average strategic complementarity weighted by firms' share of total production in the sample.

Table 2: Size of Firms' Nowcast Errors

		Size of nowcast errors[a]	
	Observations	*Industry inflation*	*Aggregate inflation*
Industry	*(1)*	*(2)*	*(3)*
Construction and Transportation	52	0.75	3.95
		(0.54)	(1.95)
Manufacturing	363	1.43	2.55
		(1.72)	(2.04)
Professional Financial Services	352	1.51	4.23
		(1.59)	(1.73)
Trade	302	0.63	2.31
		(0.90)	(1.93)
Total	1,069	1.20	3.11
		(1.49)	(2.09)

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 3: Subjective Uncertainty in Forecasts of Firms

		Subjective uncertainty (std. dev.) in forecasts[a]	
	Observations	*Industry inflation*	*Aggregate inflation*
Industry	*(1)*	*(2)*	*(3)*
Construction and Transportation	289	0.99	1.17
		(0.87)	(0.76)
Manufacturing	715	0.83	1.01
		(0.60)	(0.65)
Professional Financial Services	617	0.81	1.01
		(0.61)	(0.71)
Trade	419	0.85	1.02
		(0.63)	(0.71)
Total	2,040	0.86	1.04
		(0.66)	(0.70)

The table reports standard deviations of firms' reported distribution for their forecasts of industry and aggregate inflation. Forecasts were for yearly inflation for the 12 months ending in July 2017.

Table 4: Subjective uncertainty of firms given their number of competitors.

	Subjective uncertainty about					
	Aggregate inflation[a]		*Industry inflation*[b]		*Industry inflation rel. to aggregate inflation*[c]	
firm characteristics	(1)	(2)	(3)	(4)	(5)	(6)
Number of competitors	-0.021***	-0.024***	-0.010***	-0.007***	0.012***	0.017***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.004)
Firm controls and industry fixed effects	No	Yes	No	Yes	No	Yes
Observations	2,040	1,910	2,040	1,910	2,039	1,909
R-squared	0.036	0.050	0.009	0.020	0.007	0.017

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The table reports the result of regressing the standard deviation of firms' reported distribution for their forecast of aggregate inflation (a), and industry price change (b) on number of competitors and a set of firm controls. Columns (5) and (6) report the results of regressing the difference of the two standard deviations on the number of competitors.

Table 5: Calibration

Parameter	*Description*	*Value*	*Moment Matched*
λ	Capacity of processing information	0.7	Persistence of now-cast errors
K	Number of firms within industries	5	Average Number of Competitors
η	Elasticity of substitution within industries	6	Average Markup
ξ	Curvature of the elasticity of demand	40	Average strategic complementarity
ρ	Persistence of the growth of nominal demand	0.5	Nominal GDP growth in NZ

The table reports the calibrated values of the parameters for the dynamic model of Section 4.

Appendix

A Static Beauty Contests with Finite Players

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 A.1. Subsection A.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 A.3 proves the optimality of linear pricing strategies given Gaussian signals, and Subsection A.4 proves that when the set of available signals is rich all firms prefer to see a *single* signal. Subsection A.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 A.6 shows that the equilibrium is unique given this equivalence relationship. Subsection A.7 derives an intuitive reinterpretation of a firm's attention problem that is discussed in Section 2. Subsection A.8 contains the proofs of Propositions 1, 2, and 3 as well as the proof for Corollary 1.

A.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.[40] 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

$$\mathcal{I}(X; Y) = \frac{1}{2} \log_2 (\frac{\det(\Sigma_X)}{\det(\Sigma_{X|Y})}), \tag{10}$$

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

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

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.

Lemma 1. *Let X, Y and Z be three random variables such that $X \perp Z|Y$.*[41] *Then,*

$$\mathcal{I}(X;Y) \geq \mathcal{I}(X;Z).$$

Proof. By the chain rule for mutual information[42]

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

Notice that since $X \perp Z|Y$, then $\mathcal{I}(X;Z|Y) = 0$. Thus,

$$\mathcal{I}(X;Y) = \mathcal{I}(X;Z) + \underbrace{\mathcal{I}(X;Y|Z)}_{\geq 0} \geq \mathcal{I}(X;Z).$$

□

A.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 many independent sources of randomness 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.

[41] This forms a Markov chain: $X \rightarrow Y \rightarrow Z$.

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

Lemma 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, $S \sim \mathcal{N}(0, \Sigma) = G$. □

Before proceeding to characterization of the equilibrium, I prove the following Corollary which is going to be useful in solving the firm's problems.

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 2. *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 $\hat{S} = (S, \hat{s}) \sim \hat{G}$.*

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.[43] Let $\hat{s} \equiv \mathbf{y}'\Sigma^{-1}S + \begin{bmatrix} \sqrt{x - \mathbf{y}'\Sigma^{-1}\mathbf{y}} \\ \mathbf{0}_{N\times 1} \end{bmatrix} e_{N+1}$. Notice that $\hat{s} \in \mathcal{S}$ as it is a finite linear combination of the elements of $\mathcal{B}$. Notice that $cov(\hat{s}, S) = \mathbf{y}$ and $var(\hat{s}) = x$. Hence, $(\hat{s}, S) \sim \mathcal{N}(0, \hat{\Sigma})$. □

A.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 $p_{j,k} : S_{j,k} \to \mathbb{R}$ that maps their signal to a price. Thus, the set of firm j, k's pure strategies is

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

Moreover, the set of pure strategies for the game is

$$\mathcal{A} = \{\varsigma | \varsigma = (\varsigma_{j,k})_{j,k\in J\times K}, \varsigma_{j,k} \in \mathcal{A}_{j,k}, \forall j, k \in J \times K\}.$$

First, I show that in any equilibrium it has to be the case that firms play linear pricing strategies, meaning that $p_{j,k} = M'_{j,k}S_{j,k}$, for some $M_{j,k} \in \mathbb{R}^{n_{j,k}}$.

Lemma 3. *Take a strategy $\varsigma = (S_{j,k}, p_{j,k})_{j,k\in J\times K} \in \mathcal{A}$. Then if ς 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 ς to be an equilibrium is if given $(S_{j,k})_{j,k\in J\times K}$ under ς, $\forall j, k \in J \times K$, $p_{j,k}$ 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:

$$\begin{aligned} p^*_{j,k}(S_{j,k}) &= (1-\alpha)\mathbb{E}[q|S_{j,k}] + \alpha\frac{1}{K-1}\sum_{l\neq k}\mathbb{E}[p^*_{j,l}(S_{j,l})|S_{j,k}] \\ &= (1-\tilde{\alpha})\mathbb{E}[q|S_{j,k}] + \tilde{\alpha}\mathbb{E}[p^*_j(S_j)|S_{j,k}] \end{aligned}$$

[43] In fact, there are countably many elements in $\mathcal{B}$ that are orthogonal to S.

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

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

where $\mathbb{E}^{(0)}_{j,k}[q] \equiv \mathbb{E}[q|S_{j,k}]$ is firm j,k's expectation of the fundamental, and $\forall m \geq 1$, $\mathbb{E}^{(m)}_{j,k}[q] = K^{-1} \sum_{l \in K} \mathbb{E}[\mathbb{E}^{(m-1)}_{j,l}[q]|S_{j,k}]$ is firm j,k's m'th order higher order belief of its industry's average expectation of the fundamental. Similarly $\mathbb{E}^{(M+1)}_{j,k}[p^*_j(S_j)]$ is firm j,k's $M+1$'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:[44]

$$p^*_{j,k}(S_{j,k}) = (1-\tilde{\alpha}) \sum_{m=0}^{\infty} \tilde{\alpha}^m \mathbb{E}^{(m)}_{j,k}[q]. \tag{11}$$

Now, I just need to show that $\mathbb{E}^{(m)}_{j,k}[q]$ 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 $\Sigma_{q,S_{j,k}} \equiv cov(S_{j,k}, q) = \mathbb{E}[qS'_{j,k}]$. Also given j,k, $\forall l \neq k$, $\Sigma_{S_{j,l},S_{j,k}} = cov(S_{j,k}, S_{j,l}) = \mathbb{E}[S_{j,l}S'_{j,k}]$ and $\Sigma_{S_{j,k}} = var(S_{j,k}) = \mathbb{E}[S_{j,k}S'_{j,k}]$.

The proof for linearity of higher order expectations is by induction: notice that for $m = 0$

$$\mathbb{E}^{(0)}_{j,k}[q] = \mathbb{E}[q|S_{j,k}] = \Sigma_{q,S_{j,k}} \Sigma^{-1}_{S_{j,k}} S_{j,k},$$

which implies 0'th order expectations of firms are linear in their signals. Now suppose $\forall j,l$ $\mathbb{E}^{(m)}_{j,l}[q] = A_{j,l}(m)' S_{j,l}$ for some $A_{j,l}(m) \in \mathbb{R}^{n_{j,l}}$. Now,

$$\begin{aligned} \mathbb{E}^{(m+1)}_{j,k}[q] &= K^{-1} \sum_{l \in K} A_{j,l}(m)' \mathbb{E}[S_{j,l}|S_{j,k}] \\ &= \underbrace{K^{-1}(A_{j,l}(m) + \sum_{l \neq k}^{\prime} A_{j,l}(m) \Sigma_{S_{j,l},S_{j,k}} \Sigma^{-1}_{S_{j,k}})}_{\equiv A^{(m+1)}_{j,k} \in \mathbb{R}^{n_{j,k}}} S_{j,k}. \end{aligned}$$

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. □

[44]if 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.

Corollary 3. *If* $\varsigma = (S_{j,k} \in \mathcal{S}^{n_{j,k}}, p_{j,k}(S_{j,k}) = M'_{j,k}S_{j,k})_{j,k \in J\times K} \in \mathcal{A}$ *is an equilibrium, then* $\forall j,k \in J \times K$,

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

Proof. From the proof of Lemma 3 that if ς 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,

$$M'_{j,k}S_{j,k} = (1-\alpha)\Sigma_{q,S_{j,k}}\Sigma^{-1}_{S_{j,k}}S_{j,k} + \alpha\frac{1}{K-1}\sum_{l\neq k}M'_j\Sigma_{S_{j,l},S_{j,k}}\Sigma^{-1}_{S_{j,k}}S_{j,k},$$

or simply

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

□

Given the results in this section, I restrict the set of strategies to those with linear pricing schemes that sarisfy Corollary 3:

$$\mathcal{A}^* = \{\varsigma \in \mathcal{A} | \varsigma \text{ satisfies Corollary 3}\}.$$

A.4 The Attention Problem of Firms

Take a strategy $\varsigma \in \mathcal{A}^*$ such that

$$\varsigma = (S_{j,k} \in \mathcal{S}^{n_{j,k}}, p_{j,k} = M'_{j,k}S_{j,k})_{j,k\in J\times K}.$$

For ease of notation let $p(\varsigma_{j,k}) \equiv M'_{j,k}S_{j,k}$, $\forall j,k \in J \times K$. Also, let $\varsigma_{-(j,k)} \equiv \varsigma \backslash \varsigma_{j,k}$. Moreover, for any given firm $j,k \in J\times K$, let

$$\theta_{j,k}(\varsigma_{-(j,k)}) \equiv (q, (p(\varsigma_{j,l}))_{l\neq k}, (p(\varsigma_{m,n}))_{m\neq j, n\in K})'$$

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

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

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

$$\begin{aligned} \min_{\hat{\varsigma}_{j,k} \in \mathcal{A}_{j,k}} & L_{j,k}(\hat{\varsigma}_{j,k}, \varsigma_{-(j,k)}) \equiv \mathbb{E}[(p(\hat{\varsigma}_{j,k}) - \mathbf{w}'\theta_{j,k}(\varsigma_{-(j,k)}))^2 | S(\hat{\varsigma}_{j,k})] \\ s.t. \quad & \mathcal{I}(S(\hat{\varsigma}_{j,k}); \theta_{j,k}(\varsigma_{-(j,k)})) \leq \kappa, \end{aligned} \tag{12}$$

where given the joint distribution of $(S(\hat{\varsigma}_{j,k}), \theta_{j,k}(\varsigma_{-(j,k)}))$, the mutual information is defined by Equation (10) in Section A.1. It is also useful to restate the definition of the equilibrium given this notation:

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

$$\begin{aligned} \varsigma_{j,k} \quad = \quad & \text{argmin}_{\varsigma'_{j,k} \in \mathcal{A}_{j,k}} L_{j,k}(\varsigma'_{j,k}, \varsigma_{-(j,k)}) \\ s.t. \quad & \mathcal{I}(S(\varsigma_{j,k}); \theta_{j,k}(\varsigma_{-(j,k)})) \leq \kappa. \end{aligned} \tag{13}$$

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 $\varsigma \in \mathcal{A}^*$, and show that it is an equivalence.

Definition. For any two distinct elements $\{\varsigma^1_{j,k}, \varsigma^2_{j,k}\} \subset \mathcal{A}_{j,k}$, and given $\varsigma = (\varsigma_{j,k}, \varsigma_{-(j,k)}) \in \mathcal{A}^*$, we say $\varsigma^1_{j,k} \sim_{j,k|\varsigma} \varsigma^2_{j,k}$ if

$$L_{j,k}(\varsigma^1_{j,k}, \varsigma_{-(j,k)}) = L_{j,k}(\varsigma^2_{j,k}, \varsigma_{-(j,k)}),$$

where $L_{j,k}(.,.)$ is defined as in Equation (12).

Lemma 4. *$\forall j,k \in J \times K$ and $\forall \varsigma \in \mathcal{A}^*$, $\sim_{j,k|\varsigma}$ is an equivalence relation.*

Proof. Reflexivity, symmetry and transitivity are trivially satisfied by properties of equality. □

By definition, notice that the agent is indifferent between elements of an equivalence class. Now, given $\varsigma = (\varsigma_{j,k}, \varsigma_{-(j,k)}) \in \mathcal{A}^*$, let $[\hat{\varsigma}_{j,k}]_\varsigma \equiv \{\varsigma'_{j,k} \in \mathcal{A}_{j,k} | \varsigma'_{j,k} \sim_{j,k|\varsigma} \hat{\varsigma}_{j,k}\}$. The following lemma shows there is always a deviation with a single dimensional signal that requires less attention.

Lemma 5. *For any* $j,k \in J \times K$, $\forall \varsigma = (\varsigma_{j,k}, \varsigma_{-(j,k)}) \in \mathcal{A}^*$, $\exists \hat{\varsigma}_{j,k} \in [\varsigma_{j,k}]_\varsigma$ *such that the agent observes only one signal under* $\hat{\varsigma}_{j,k}$ *and*

$$\mathcal{I}(S(\hat{\varsigma}_{j,k}); \theta_{j,k}(\varsigma_{-(j,k)})) \le \mathcal{I}(S(\varsigma_{j,k}); \theta_{j,k}(\varsigma_{-(j,k)})).$$

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 $\Sigma_{\varsigma_{j,k}} \equiv var(S(\varsigma_{j,k}))$, $\Sigma_{\theta_{j,k},\varsigma_{j,k}} \equiv cov(\theta_{j,k}(\varsigma_{-(j,k)}), S(\varsigma_{j,k}))$ and $\Sigma_{\theta_{j,k}} \equiv var(\theta_{j,k}(\varsigma_{-(j,k)}))$. 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 3

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

Notice that

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

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

$$\begin{aligned} L_{j,k}(\hat{\varsigma}_{j,k}, \varsigma_{-(j,k)}) &= \mathbf{w}'var(\theta_{j,k}(\varsigma_{-(j,k)})|\hat{s}_{j,k})\mathbf{w} \\ &= \mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w} - \mathbf{w}'\Sigma_{\theta_{j,k},\varsigma_{j,k}}\Sigma^{-1}_{\varsigma_{j,k}}\Sigma'_{\theta_{j,k},\varsigma_{j,k}}\mathbf{w} \\ &= L_{j,k}(\varsigma_{j,k}, \varsigma_{-(j,k)}). \end{aligned}$$

Thus, $\hat{\varsigma}_{j,k} \in [\varsigma_{j,k}]_\varsigma$. Also, observe that $\theta_{j,k}(\varsigma_{-(j,k)}) \perp \hat{s}_{j,k}|S(\varsigma_{j,k})$. Therefore, by the data processing inequality in Lemma 1

$$\mathcal{I}(\hat{s}_{j,k}; \theta_{j,k}(\varsigma_{-(j,k)})) \le \mathcal{I}(S(\varsigma_{j,k}); \theta_{j,k}(\varsigma_{-(j,k)})).$$

Finally, observe that $p_{j,k}(\hat{\varsigma}_{j,k}, \varsigma_{-(j,k)}) = p_{j,k}(\varsigma) = \mathbf{w}'\Sigma_{\theta_{j,k},\varsigma_{j,k}}\Sigma^{-1}_{\varsigma_{j,k}}S(\varsigma_{j,k})$. 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 $\hat{\varsigma}_{j,k}$. □

A.5 Equilibrium Signals

Let $\mathcal{E}$ denote the set of equilibria for the game:

$$\mathcal{E} = \{\varsigma \in \mathcal{A} | \varsigma \text{ is an equilibrium as stated in Statement (13)}\}.$$

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 6. *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 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 $\varsigma^{J,K} \in \mathcal{A}^{**}$. □

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 7. *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.[45]

Now, from Lemma 5 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 2 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 $s^1_{j,k}$ and $s^2_{j,k}$ in $\mathcal{S}$, $(s^1_{j,k}, \theta_{j,k}(\varsigma_{-(j,k)}))$ and $(s^2_{j,k}, \theta_{j,k}(\varsigma_{-(j,k)}))$ have the same joint distribution. Then, $var(\theta_{j,k}(\varsigma_{-(j,k)})|s^1_{j,k}) = var(\theta_{j,k}(\varsigma_{-(j,k)})|s^2_{j,k})$ which implies that

$$L_{j,k}((s^1_{j,k}, 1), \varsigma_{-(j,k)}) = L_{j,k}((s^2_{j,k}, 1), \varsigma_{-(j,k)}).$$

Moreover,

$$\begin{aligned} \mathcal{I}(s^1_{j,k}; \theta_{j,k}(\varsigma_{-(j,k)})) &= \frac{1}{2}\log_2(\frac{\det(var(\theta_{j,k}(\varsigma_{-(j,k)}))}{\det(var(\theta_{j,k}(\varsigma_{-(j,k)})|s^1_{j,k})}) \\ &= \frac{1}{2}\log_2(\frac{\det(var(\theta_{j,k}(\varsigma_{-(j,k)}))}{\det(var(\theta_{j,k}(\varsigma_{-(j,k)})|s^2_{j,k})}) \\ &= \mathcal{I}(s^1_{j,k}; \theta_{j,k}(\varsigma_{-(j,k)})). \end{aligned}$$

Therefore, the firm is indifferent between $s^1_{j,k}$ and $s^2_{j,k}$. □

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.

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

Lemma 8. *Suppose* $\varsigma = (s^*_{j,k} \in \mathcal{S}, 1) \in \mathcal{A}^{**}$ *is an equilibrium, then* $\forall j,k \in J \times K$,

$$s^*_{j,k} = \lambda \mathbf{w}'\theta_{j,k}(\varsigma_{-(j,k)}) + z_{j,k},\ z_{j,k} \perp \theta_{j,k}(\varsigma_{-(j,k)}),\ var(z_{j,k}) = \lambda(1-\lambda)var(\mathbf{w}'\theta_{j,k}(\varsigma_{-(j,k)})).$$

Proof. For firm $j,k \in J \times K$, let $\Sigma_{\theta_{j,k}}$ denote the covariance matrix of $\theta_{j,k}(\varsigma_{-(j,k)})$. From Lemma 5 it is sufficient to look at deviations of the form $(s_{j,k} \in \mathcal{S}, 1)$. For a given $s_{j,k} \in \mathcal{S}$, $(s_{j,k}, \theta_{j,k}(\varsigma_{-(j,k)})) \sim \mathcal{N}(0, \Sigma_{s_{j,k},\theta_{j,k}})$, 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 of all, recall that for $(s_{j,k} \in \mathcal{S}, 1)$ to be optimal, it has to be the case that

$$\begin{aligned} p_{j,k} &= 1 \times s_{j,k} \\ &= \mathbf{w}'\mathbb{E}[\theta_{j,k}(\varsigma_{-(j,k)})|s_{j,k}] \\ &= x^{-2}\mathbf{w}'\mathbf{y}s_{j,k}. \end{aligned}$$

Thus,

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

Now, given $s_{j,k} \in \mathcal{S}$, the firm's loss in profits is

$$var(\mathbf{w}'\theta_{j,k}(\varsigma_{-(j,k)})|s_{j,k}) = \mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w} - x^{-2}(\mathbf{w}'\mathbf{y})^2$$

and the capacity constraint

$$\begin{aligned} & \mathcal{I}(s_{j,k}; \theta_{j,k}(\varsigma_{-(j,k)})) \leq \kappa \\ \Leftrightarrow\ & \frac{1}{2}\log_2(|\mathbf{I} - x^{-2}\Sigma^{-1}_{\theta_{j,k}}\mathbf{y}\mathbf{y}'|) \geq -\kappa \\ \Leftrightarrow\ & x^{-2}\mathbf{y}'\Sigma^{-1}_{\theta_{j,k}}\mathbf{y} \leq \lambda \equiv 1 - 2^{-2\kappa}. \end{aligned}$$

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

$$\begin{aligned} & \min_{(x,\mathbf{y})} \mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w} - x^{-2}(\mathbf{w}'\mathbf{y})^2 \\ s.t.\ \ & x^{-2}\mathbf{y}'\Sigma^{-1}_{\theta_{j,k}}\mathbf{y} \leq \lambda \end{aligned}$$

The solution can be derived by taking first order conditions, but there is simpler a way. Notice that by Cauchy-Schwarz inequality

$$
\begin{aligned}
x^{-2}(\mathbf{w}'\mathbf{y})^2 &= x^{-2}(\Sigma_{\theta_{j,k}}^{\frac{1}{2}}\mathbf{w})'(\Sigma_{\theta_{j,k}}^{-\frac{1}{2}}\mathbf{y}) \\
&\leq x^{-2}(\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w})(\mathbf{y}'\Sigma_{\theta_{j,k}}^{-1}\mathbf{y}).
\end{aligned}
$$

Therefore,

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

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

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

notice that $cov(s_{j,k}^*, \theta_{j,k}(\varsigma_{-(j,k)})) = \lambda\Sigma_{\theta_{j,k}}\mathbf{w}$, and $var(s_{j,k}^*) = \lambda\mathbf{w}'\Sigma_{\theta_{j,k}}\mathbf{w}$. 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 $var(z_{j,t}) = \lambda(1-\lambda)var((1-\alpha)q + \alpha\frac{1}{K-1}\sum_{l\neq k} s^*_{j,l})$. □

A.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 9. *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

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

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

$$\begin{aligned} cov(p_j, q) &= \lambda(1-\alpha)\mathbf{1}\underbrace{cov(q,q)}_{=1} + \lambda\alpha\gamma(\mathbf{1}\mathbf{1}' - \mathbf{I})cov(p_j,q) + \underbrace{cov(z_j,q)}_{=0} \\ &\Rightarrow \\ ((1+\lambda\alpha\gamma)\mathbf{I} - \lambda\alpha\gamma\mathbf{1}\mathbf{1}')cov(p_j,q) &= \lambda(1-\alpha)\mathbf{1} \end{aligned}$$

Now, notice that $(1+\lambda\alpha\gamma)\mathbf{I} - \lambda\alpha\gamma\mathbf{11'}$ is a symmetric matrix whose diagonal elements are strictly larger than its off diagonal elements. It is straight forward to show that

$$((1+\lambda\alpha\gamma)\mathbf{I} - \lambda\alpha\gamma\mathbf{11'})^{-1} = \frac{1}{1+\alpha\lambda\gamma}\mathbf{I} + \frac{\alpha\lambda\gamma}{(1+\alpha\lambda\gamma)(1-\alpha\lambda)}\mathbf{11'}.$$

Since $\alpha\lambda < 1 \Rightarrow 1-\alpha\lambda > 0$ and this inverse exists. Now, we have

$$\begin{aligned} cov(p_j, q) &= (\frac{1}{1+\alpha\lambda\gamma}\mathbf{I} + \frac{\alpha\lambda\gamma}{(1+\alpha\lambda\gamma)(1-\alpha\lambda)}\mathbf{11'})\lambda(1-\alpha)\mathbf{1} \\ &= \frac{\lambda - \lambda\alpha}{1-\lambda\alpha}\mathbf{1}. \end{aligned}$$

Thus, in any equilibrium, the covariance of any firm's price with the fundamental q has be to equal to $\delta \equiv \frac{\lambda-\lambda\alpha}{1-\lambda\alpha}$.

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

$$cov(p_j, p_{l,m}) = \lambda(1-\alpha)\mathbf{1}\underbrace{cov(q, p_{l,m})}_{=\delta} + \lambda\alpha\gamma(\mathbf{11'} - \mathbf{I})cov(p_j, p_{l,m}) + \underbrace{cov(z_j, p_{l,m})}_{=0}.$$

With a similar method as above, 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 simply 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

$$\begin{aligned} p_j &= \lambda(1-\alpha)\mathbf{1}q + \lambda\alpha\gamma(\mathbf{11'} - \mathbf{I})p_j + z_j \\ &= \mathbf{B}(\lambda(1-\alpha)\mathbf{1}q + z_j) \end{aligned}$$

where $\mathbf{B} \equiv \frac{1}{1+\alpha\lambda\gamma}\mathbf{I} + \frac{\alpha\lambda\gamma}{(1+\alpha\lambda\gamma)(1-\alpha\lambda)}\mathbf{11'}$. 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 $\mathbf{w}_k' p_j = \sum_{l\neq k} p_{j,l}$. Let $\mathbf{e}_k$ be the k'th column of the identity matrix. Hence,

$$\begin{aligned} \mathbf{e}_k'\mathbf{D}_j\mathbf{e}_k &= \lambda(1-\lambda)(1-\alpha)(1-\alpha+2\alpha\delta) + \lambda(1-\lambda)\alpha^2\gamma^2\mathbf{w}_k'(\delta^2\mathbf{1}\mathbf{1}' + \mathbf{B}\Sigma_{z,j}\mathbf{B})\mathbf{w}_k \\ &= (\lambda^{-1}-1)\delta^2 + \lambda(1-\lambda)\alpha^2\gamma^2\mathbf{w}_k'\mathbf{D}_j\mathbf{B}\mathbf{w}_k \end{aligned}$$

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

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

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

A.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

$$\begin{aligned} \rho_q(s_{j,k}) &\equiv cor(s_{j,k}, q), \\ \rho_v(s_{j,k}) &\equiv cor(s_{j,k}, v_{j,-k}). \end{aligned}$$

Notice that firm j,k's loss in profit given that they observe $s_{j,k}$ is

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

Observe that

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

now, to derive the information constraint in terms of the two correlations:

$$\mathcal{I}(s_{j,k}; (q, p^*_{j,-k})) \leq \kappa \Leftrightarrow \frac{1}{2}\log_2(\frac{var(s_j)}{var(s_{j,k}|(q, p^*_{j,-k}))}) \leq \kappa$$

Now, notice that

$$\begin{aligned} \frac{var(s_j|(q, p^*_{j,-k}))}{var(s_j)} &= 1 - [\ \rho_q(s_j) \quad \delta\rho_q(s_j) + \sigma_v\rho_v(s_j)] \begin{bmatrix} 1 & \delta \\ \delta & \delta^2 + \sigma_v^2 \end{bmatrix}^{-1} \begin{bmatrix} \rho_q(s_j) \\ \delta\rho_q(s_j) + \sigma_v^2 \end{bmatrix} \\ &= 1 - (\rho_q(s_j)^2 + \rho_v(s_j)^2). \end{aligned}$$

Thus, the information constraint becomes

$$\frac{1}{2}\log_2(\frac{1}{1-(\rho_q^2(s_j)+\rho_v^2(s_j))}) \leq \kappa \Leftrightarrow \rho_q^2(s_j)+\rho_v^2(s_j) \leq \lambda \equiv 1-2^{-2\kappa}.$$

So j,k's problem reduces to

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

Since 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. So instead of choosing signals, the agent can choose the correlations:

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

A.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 B.

In the order of appearance in the text.

Proof of Proposition 1.

Part 1. Given the result in Lemma 9, notice that since attention is strictly increasing in the squared correlation:

$$\begin{aligned} \rho_q^{*2} &= cor(p_{j,k},q)^2 \\ &= \frac{cov(p_{j,k},q)^2}{var(p_{j,k})} \\ &= \frac{K-1+\alpha\delta}{K-1+\alpha\lambda}\lambda. \end{aligned}$$

but 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.

Part 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 $\frac{\partial\delta}{\partial\alpha} = \delta\frac{\lambda-1}{(1-\alpha)(1-\alpha\lambda)} < 0$.

Part 3. Shown in the proof of Lemma 9.

Proof of Proposition 2.

First of all notice that the aggregate price is given by

$$p \equiv J^{-1}K^{-1}\sum_{j,k\in J\times K} p_{j,k} = \delta q + \frac{1}{JK}\sum_{j,k\in J\times K} v_{j,k}$$

but notice that since J is large and $v_{j,k}$'s are independent across industries, the average converges to zero by law of large numbers as $J\to\infty$. Therefore,

$$p = \delta q.$$

Moreover, notice that

$$\begin{aligned}
\mathbb{E}^{j,k}[p_{j,-k}] &= \frac{cov(s_{j,k},p_{j,-k})}{var(p_{j,k})}s_{j,k} \\
&= \tilde{\lambda}p_{j,k} \\
\mathbb{E}^{j,k}[p] &= \frac{cov(s_{j,k},p)}{var(p_{j,k})}p_{j,k} \\
&= \frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda}\lambda p_{j,k}
\end{aligned}$$

where $\tilde{\lambda} = \frac{\lambda(K-1)+\alpha\lambda}{K-1+\alpha\lambda} > \lambda$ is defined as in the proof of Lemma 9. So, $\overline{\mathbb{E}^{j,k}[p_{j,-k}]} = \tilde{\lambda}p$, $\overline{\mathbb{E}^{j,k}[p]} = \frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda}\lambda p$. notice that

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

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

Proof of Corollary 1.

Conditional on realization of the aggregate price

$$\begin{aligned} |p - \overline{\mathbb{E}^{j,k}[p]}| &= (1 - \frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda}\lambda)|p| \\ &> (1-\tilde{\lambda})|p| \\ &= |p - \overline{\mathbb{E}^{j,k}[p_{j,-k}]}|. \end{aligned}$$

Proof of Proposition 3.

Since *knowledge* is directly related to mutual information (as defined in Definition 3), and mutual information in this static setting reduces to correlations, we need to show

$$cor(p_{j,k}, p_{j,-k}) \geq cor(p_{j,k}, q) = cor(p_{j,k}, p).$$

By plugging in the unique equilibrium distribution from the proof of Lemma 9, we get this holds if and only if

$$\frac{1-\alpha\tilde{\lambda}}{1-\alpha\lambda}\lambda \leq \frac{(K-1)\tilde{\lambda}^2}{1+(K-2)\tilde{\lambda}}.$$

Moving the terms around, this can be rearranged to

$$\alpha\tilde{\lambda} \geq \frac{1}{2},$$

meaning that the necessary and sufficient condition for the result is when this inequality holds. Now, notice that if $\alpha\lambda \geq \frac{1}{2}$, since $\tilde{\lambda} \geq \lambda$, then $\alpha\tilde{\lambda} \geq \frac{1}{2}$. Hence, $\alpha\lambda \geq \frac{1}{2}$ is a sufficient condition.

B Online Appendix

The Appendix is organized as follows. Subsection B.1 extends the set of available information defined in Appendix A.2 to the dynamic environment. Subsection B.2 includes all the derivations for the dynamic model that are omitted in the main text. Subsection B.3 discusses the degree of strategic complementarity implied by the Kimball aggregator. Subsection B.4 contains the proofs of the Propositions 4, 5, 6 and 7. Subsection B.5 discusses the computational method that I use for solving the dynamic model.

B.1 Available Information in the Dynamic Model

The set of available signals in the dynamic model is an extension of the one defined in Appendix A.2. The main difference is the notion of time and the fact that at every period nature draws new shocks and the set of 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 A.2, let $\mathcal{S}_t$ be the set of all finite linear combinations of these uncorrelated noises. Now, define

$$\mathcal{S}^t = \{\sum_{s=0}^{\infty} a_\tau e_{t-\tau} | \forall \tau \geq 0, a_\tau \in \mathbb{R}, e_{t-\tau} \in \mathcal{S}_{t-\tau}\}, \forall t.$$

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 fundamental in the economy.

B.2 Derivations

Solution to Household's Problem (5).

Let $\beta^t \varphi_{1,t}$ and $\beta^t \varphi_{2,t}$ be the Lagrange multipliers on household's budget and aggregation constraints, respectively.

For ease of notation let $\mathcal{C}_{j,t} \equiv (C_{j,1,t}, \ldots, C_{j,K,t})$ be the vector of household's consumption from firms in industry $j \in J$, so that $C_{j,t} \equiv \Phi(\mathcal{C}_{j,t})$. 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

$$P_{j,k,t} = \frac{1}{J}\frac{\varphi_{2,t}}{\varphi_{1,t}} C_t \frac{\Phi_k(\mathcal{C}_{j,t})}{\Phi(\mathcal{C}_{j,t})} \tag{14}$$

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

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

where the equality under curly bracket is from Euler theorem for homogeneous functions as $\Phi(.)$ 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 (14) notice that

$$\mathcal{P}_{j,t} \equiv (P_{j,1,t}, \ldots, P_{j,K,t}) = \nabla \log(\Phi(\frac{\mathcal{C}_{j,t}}{J^{-1} P_t C_t})).$$

Now, I need to show that this function is invertible to prove that a demand function exists. For ease of notation, define function $f : \mathbb{R}^K \to \mathbb{R}^K$ such that $f(\mathbf{x}) \equiv \nabla \log(\Phi(\mathbf{x}))$. Notice that $f(.)$ is homogeneous of degree -1, and the m, n'th element of its Jacobian, denoted by matrix $\mathcal{J}^f(\mathbf{x})$, is given by

$$\mathcal{J}^f_{m,n}(\mathbf{x}) \equiv \frac{\partial}{\partial x_n} \frac{\Phi_m(\mathbf{x})}{\Phi(\mathbf{x})} = \frac{\Phi_{m,n}(\mathbf{x})}{\Phi(\mathbf{x})} - \frac{\Phi_n(\mathbf{x})}{\Phi(\mathbf{x})}\frac{\Phi_m(\mathbf{x})}{\Phi(\mathbf{x})}.$$

Let $\mathbf{1}$ be the unit vector in $\mathbb{R}^K$. Since $\Phi(.)$ is symmetric along its arguments, for any $k \in (1, \ldots, K)$, $\Phi_1(\mathbf{1}) = \Phi_k(\mathbf{1})$, $\Phi_{11}(\mathbf{1}) = \Phi_{kk}(\mathbf{1}) < 0$. Since $\Phi(.)$ is homogeneous of degree 1, by Euler's theorem we have

$$\Phi(\mathbf{1}) = \sum_{k\in K} \Phi_k(\mathbf{1}) = K\Phi_1(\mathbf{1}).$$

Also, since $\Phi_k(.)$ is homogeneous of degree zero[46], similarly we have

$$0 = 0 \times \Phi_k(\mathbf{1}) = \sum_{l\in K} \Phi_{kl}(\mathbf{1}).$$

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

So, for any $l \neq k$,

$$\Phi_{kl}(\mathbf{1}) = -\frac{1}{K-1}\Phi_{11}(\mathbf{1}) > 0.$$

This last equation implies that $\mathcal{J}^f(\mathbf{1})$ is an invertible matrix.[47] Therefore, by inverse function theorem $f(.)$ is invertible in an open neighborhood around $\mathbf{1}$, and therefore any symmetric point $\mathbf{x} = x.\mathbf{1}$ such that $x > 1$. We can write

$$\frac{\mathcal{C}_{j,t}}{J^{-1}P_tC_t} = f^{-1}(\mathcal{P}_{j,t}).$$

It is straight forward to show that $f^{-1}(.)$ is homogeneous of degree -1 simply because $f(\mathbf{x})$ is homogeneous of degree -1: for any $\mathbf{x} \in \mathbb{R}^K$,

$$\begin{aligned} f^{-1}(a\mathbf{x}) &= f^{-1}(af(f^{-1}(\mathbf{x})) \\ &= f^{-1}(f(a^{-1}f^{-1}(\mathbf{x})) \\ &= a^{-1}f^{-1}(\mathbf{x}). \end{aligned}$$

Now, notice that

$$C_{j,k,t} = J^{-1}P_tC_tf_k^{-1}(\mathcal{P}_{j,t}),$$

where $f_k^{-1}(\mathbf{x})$ is the k'th element of the vector $f^{-1}(\mathcal{P}_{j,t})$. Finally, since $f(.)$ is symmetric across its arguments, so is $f^{-1}(\mathcal{P}_{j,t})$, meaning that $f_k^{-1}(\mathcal{P}_{j,t}) = f_1^{-1}(\sigma_{k,1}(\mathcal{P}_{j,t}))$, where $\sigma_{k,1}(\mathcal{P}_{j,t})$ is a permutation that changes the places of the first and k'th element of the vector $\mathcal{P}_{j,t}$. Now, to get the notation in the text let $(P_{j,k,t}, P_{j,-k,t}) \equiv \sigma_{k,1}(\mathcal{P}_{j,t})$ and $\mathcal{D}(\mathbf{x}) \equiv J^{-1}f_1^{-1}(\mathbf{x})$, which gives us the notation in the text:

$$C_{j,k,t} = P_tC_t\mathcal{D}(P_{j,k,t}, P_{j,-k,t}),$$

where $\mathcal{D}(.,.)$ is homogeneous of degree -1.

Finally, the optimality conditions of the household's problem with respect to B_t, C_t and

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

L_t are

$$\begin{aligned} w.r.t.\, C_t: &\qquad C_t^{-1} = \varphi_{2,t} = P_t\varphi_{1,t} \\ w.r.t.\, B_t: &\qquad \varphi_{1,t} = \beta(1+i_t)\mathbb{E}_t^f[\varphi_{1,t+1}] \\ w.r.t.\, L_t: &\quad L_t = \begin{cases} 0 & \phi > \varphi_{1,t}W_t \\ L \in [0,\infty] & \phi = \varphi_{1,t}W_t \\ \infty & \phi < \varphi_{1,t}W_t \end{cases} \end{aligned}$$

The optimality condition with respect to L_t is not a first order condition because household's disutility from labor is linear. Hence, household will supply a positive but finite labor supply if and only if $\phi = \varphi_{1,t}W_t$ at which point she is indifferent in supplying any amount of labor. It is only when $\phi = \varphi_{1,t}W_t$ that household supplies finite labor.[48] Hence, the optimality conditions are

$$P_tC_t = \beta(1+i_t)\mathbb{E}_t^f[P_{t+1}C_{t+1}],\ \phi P_tC_t = W_t.$$

Loss Function of the Firms.

Let $\Pi(P_{j,k,t}, P_{j-k,t}, W_t) = (P_{j,k,t} - (1-\bar{\mathrm{s}})W_t)\mathcal{D}(P_{j,k,t}, P_{j,-k,t})$ denote the profit function of the firm following the text. Notice that this function is homogeneous of degree 1 as $\mathcal{D}(.,.)$ 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

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

Now, 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

$$P^*_{j,k,t} = \text{argmax}_x \Pi(x, P_{j,-k,t}, W_t).$$

Notice that

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

has the same solution as profit maximization problem of the firm because $L(.)$ is also homo-

[48] This is also the limit of the household's labor supply curve when the Frisch elasticity goes to infinity.

geneous 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 $\forall t$, $P_{j,k,t} = P_{j,l,t}|_{\forall l \neq k} = \bar{P}, W_t = \phi\bar{Q}$ such that

$$\bar{P} = \text{argmax}_x \Pi(x, \bar{P}, \phi).$$

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

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

But notice that $L(\bar{P}, \bar{P}, \phi\bar{Q}) = 0$, and $p^*_{j,k,t} = \frac{P^*_{j,k,t} - \bar{P}}{\bar{P}}$ is such that

$$\Pi_1(P^*_{j,k,t}, P_{j,-k,t}, \phi Q_t) = 0,$$

meaning that

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

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

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

which implies that up to this second order approximation the profit maximization of the firm is equivalent to

$$\min_{(p_{j,k,t}:S^t_{j,k}\to\mathbb{R})_{t=0}^{\infty}} \mathbb{E}[\sum_{t=0}^{\infty} \beta^t (p_{j,k,t} - p^*_{j,k,t})^2 | S^{-1}_{j,k}].$$

The following part of the Appendix derives the specific form of $p^*_{j,k,t}$.

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

$$\begin{aligned} P^*_{j,k,t} &= \frac{\varepsilon_D(P^*_{j,k,t}, P_{j,-k,t})}{\varepsilon_D(P^*_{j,k,t}, P_{j,-k,t}) - 1}(1 - \bar{\mathsf{s}})W_t \\ &= \frac{\varepsilon_D(P^*_{j,k,t}, P_{j,-k,t})}{\varepsilon_D(P^*_{j,k,t}, P_{j,-k,t}) - 1}(1 - \bar{\mathsf{s}})\phi Q_t, \end{aligned}$$

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

$$\varepsilon^{\varepsilon}_D(P_{j,k,t}, P_{j,-k,t}) \equiv \frac{P_{j,k,t}}{\varepsilon_D(P_{j,k,t}, P_{j,-k,t})} \frac{\partial}{\partial P_{j,k,t}} \varepsilon_D(P_{j,k,t}, P_{j,-k,t}).$$

Notice that since $\mathcal{D}(.,.)$ is homogeneous of degree -1, then $\varepsilon_D(.,.)$ and $\varepsilon^{\varepsilon}_D(.,.)$ are both homogeneous of degree zero. For ease of notation let $\varepsilon_D \equiv \varepsilon_D(1,1)$ and $\varepsilon^{\varepsilon}_D \equiv \varepsilon^{\varepsilon}_D(1,1)$.

Now, recall from the previous section that $p^*_{j,k,t}$ is a 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

$$\alpha \equiv \frac{\varepsilon_D^{\varepsilon}}{\varepsilon_D^{\varepsilon} + \varepsilon_D - 1}. \tag{15}$$

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.

Derivation of Demand Given Elasticities.

I assumed that every firm's own-price elasticity has the form

$$\varepsilon(P_{j,k,t}, P_{j,-k,t}) \equiv -\frac{D_1(P_{j,k,t}, P_{j,-k,t})}{D(P_{j,k,t}, P_{j,-k,t})} P_{j,k,t} = \eta - (\eta - 1) K^{\xi} \left(\frac{P_{j,k,t}^{1-\eta}}{\sum_{k \in K} P_{j,k,t}^{1-\eta}} \right)^{1+\xi}.$$

A particular solution to this partial differential equation is

$$\log(D(P_{j,k,t}, P_{j,-k,t})) = \log(\frac{P_{j,k,t}^{-\eta}}{\sum_{l \neq k} P_{j,l,t}^{1-\eta}}) - \frac{K^{\xi}}{1+\xi} \left(\frac{P_{j,k,t}^{1-\eta}}{\sum_{l \neq k} P_{j,l,t}^{1-\eta}} \right)^{1+\xi} {}_2F_1(1+\xi, 1+\xi; 2+\xi; -\frac{P_{j,k,t}^{1-\eta}}{\sum_{l \neq k} P_{j,l,t}^{1-\eta}}),$$

where ${}_2F_1(., .; .; .)$ is the hypergeometric function. This is also the particular solution to the above PDE that coincides with the CES demand when $\xi = 0$. To see this, we use the identity ${}_2F_1(1, 1; 2; x) = \frac{\log(1-x)}{x}$. Therefore, when $\xi = 0$, we have

$$\log(D(P_{j,k,t}, P_{j,-k,t})) = \log(\frac{P_{j,k,t}^{-\eta}}{\sum_{k \in K} P_{j,k,t}^{1-\eta}}).$$

Now, given the particular solution to the PDE we can define the system of equations,

$$(C_{j,k,t})_{k \in K} = J^{-1} Q_t \mathcal{D}(P_{j,k,t})_{k \in K},$$

and the inverse of this function gives us a system of first order partial differential equations in terms of the function $\Phi(.)$ as shown in derivations of household's utility function.

B.3 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 ($\alpha = 0.9$).

3. Strategic complementarity is increasing with the elasticity of substitution within industries ($\frac{\partial \alpha}{\partial \eta} \geq 0$).

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

$$1 = K^{-1} \sum_{k \in K} f(\frac{KC_{j,k,t}}{\Phi(C_{j,1,t}, \ldots, C_{j,K,t})}), \tag{16}$$

where $f(.)$ is at least thrice differentiable, and $f(1) = 1$ (so that $\Phi(1, \ldots, 1) = K$). Observe that this coincides with the CES aggregator when $f(x) = x^{\frac{\eta-1}{\eta}}$. To derive the demand functions, recall that the first order conditions of the household's problem are

$$P_{j,k,t} = J^{-1} Q_t \frac{\frac{\partial}{\partial C_{j,k,t}} C_{j,t}}{C_{j,t}}, \forall j, k$$

where $C_{j,t} = \Phi(C_{j,1,t}, \ldots, C_{j,K,t})$. Implicit differentiation of Equation (16) gives

$$P_{j,k,t} = J^{-1} Q_t \frac{f'(\frac{KC_{j,k,t}}{C_{j,t}})}{\sum_{l \in K} C_{j,l,t} f'(\frac{KC_{j,l,t}}{C_{j,t}})}, \forall j, k. \tag{17}$$

To invert these functions and get the demand for every firm in terms of their competitors' prices, guess that there exists a function $F : \mathbb{R}^K \to \mathbb{R}$ such that

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

I verify this guess by plugging in this guess to Equation (17):

$$\begin{aligned}
\forall j, l : \qquad & f'(\frac{KC_{j,l,t}}{C_{j,t}}) = P_{j,l,t} F(P_{j,1,t}, \ldots, P_{j,K,t}) \\
\Rightarrow \quad & \frac{KC_{j,l,t}}{C_{j,t}} = f'^{-1}(P_{j,l,t} F(P_{j,1,t}, \ldots, P_{j,K,t})) \\
\Rightarrow \quad & 1 = K^{-1} \sum_{l \in K} f(\frac{KC_{j,l,t}}{C_{j,t}}) = K^{-1} \sum_{l \in K} f(f'^{-1}(P_{j,l,t} F(P_{j,1,t}, \ldots, P_{j,K,t})))
\end{aligned}$$

Thus, the guess implies that the function $F(.)$ is implicitly defined by

$$1 = K^{-1} \sum_{k \in K} f(f'^{-1}(P_{j,l,t} F(P_{j,1,t}, \dots, P_{j,K,t}))),$$

which is well-defined as the expression only depends on $(P_{j,k,t})_{k \in K}$ and hence $F(.)$ only depends on the vector of these prices. This verifies my guess for existence of such a function. It is straight forward to show that $F(.)$ is symmetric across its arguments and homogeneous of degree -1.[49] Now, notice that given these derivations,

$$\begin{aligned}
J^{-1} Q_t F(P_{j,1,t}, \dots, P_{j,K,t}) &= \sum_{l \in K} C_{j,l,t} f'(\frac{K C_{j,l,t}}{C_{j,t}}) \\
&= \frac{C_{j,t}}{K} \sum_{l \in K} \frac{K C_{j,l,t}}{C_{j,t}} f'(\frac{K C_{j,l,t}}{C_{j,t}}) \\
&= \frac{C_{j,t}}{K} \sum_{l \in K} f'^{-1}(P_{j,l,t} F(P_{j,1,t}, \dots, P_{j,K,t})) P_{j,l,t} F(P_{j,1,t}, \dots, P_{j,K,t}) \\
&\Rightarrow \\
\frac{C_{j,t}}{K} &= \frac{J^{-1} Q_t}{\sum_{l \in K} P_{j,l,t} f'^{-1}(P_{j,l,t} F(P_{j,1,t}, \dots, P_{j,K,t}))}.
\end{aligned}$$

Now, notice that

$$\begin{aligned}
C_{j,k,t} &= \frac{C_{j,t}}{K} f'^{-1}(P_{j,k,t} F(P_{j,1,t}, \dots, P_{j,K,t})) \\
&= J^{-1} Q_t \frac{f'^{-1}(P_{j,k,t} F(P_{j,1,t}, \dots, P_{j,K,t}))}{\sum_{l \in K} P_{j,l,t} f'^{-1}(P_{j,l,t} F(P_{j,1,t}, \dots, P_{j,K,t}))}.
\end{aligned}$$

This gives us 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

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

To derive the degree of strategic complementarity, we just need to derive the elasticity and super elasticity of a single firm's demand around a point where all firms are charging the same price. The homogeneity of degree minus one for the firm's demand implies that these elasticities are independent of the scale of this symmetric point, so I do it around the point

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

$(1, \dots, 1) \in \mathbb{R}^K$. Without loss of generality since the function $F(.)$ is symmetric around its arguments, I do this for firm $j, 1$. Now, notice that

$$D(x, 1) = \frac{f'^{-1}(xF(x, 1, \dots, 1))}{(K-1)f'^{-1}(F(x, 1, \dots, 1)) + xf'^{-1}(xF(x, 1, \dots, 1))},$$

where $F(x, 1, \dots, 1)$ is implicitly defined by

$$K = (K-1)f(f'^{-1}(F(x, 1, \dots, 1))) + f(f'^{-1}(xF(x, 1, \dots, 1))).$$

Before deriving the strategic complementarity, 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.

Now, I define $\zeta(x) \equiv \frac{\partial \log(-\frac{\partial \log(f'(x))}{\partial \log(x)})}{\partial \log(x)}$ as the elasticity of the elasticity of $f'(x)$:

$$\zeta(x) = \frac{f'''(x)}{f''(x)}x - \frac{f''(x)}{f'(x)}x + 1.$$

For notational ease let $\zeta \equiv \zeta(1)$ and assume $\zeta \geq 0$ ($\zeta = 0$ corresponds to the case of CES aggregator). These assumptions ($\eta > 1$ and $\zeta \geq 0$ are sufficient for weak strategic complementarity, $\alpha \in [0, 1)$).While the usual approach in the literature is to assume $K \to \infty$ and look at super elasticities in this limit, a part of my main results revolve around the finiteness of the number of competitors and the fact that the degree of strategic complementarity is decreasing in K. Therefore, I derive the degree of strategic complementarity for any finite K. With some intense algebra we get

$$\alpha = \frac{\zeta(K-2) + (1-\eta^{-1})^2}{\zeta(K-2) + (1-\eta^{-1})K} \in [0, 1).$$

Notice that this imbeds the CES aggregator when $\zeta = 0$, in which case $\alpha = (1 - \eta^{-1})K^{-1}$. This generalization allows us to match a high degree of strategic complementarity by choosing a large ζ. However, this leads to a counterintuitive result where the degree of strategic complementarity decreases with the elasticity of substitution:

To derive the condition under which α is increasing in the elasticity of substitution η, notice that

$$\frac{\partial \alpha}{\partial \eta} = \frac{d(1-\eta^{-1})}{d\eta} \frac{\partial \alpha}{\partial(1-\eta^{-1})} = (\eta^{-2}) \frac{(1-\eta^{-1})^2 K - \zeta(K-2)(K - 2(1-\eta^{-1}))}{(\zeta(K-2) + (1-\eta^{-1})K)^2},$$

thus

$$\frac{\partial \alpha}{\partial \eta} \geq 0 \Leftrightarrow \zeta \leq \frac{(1-\eta^{-1})^2 K}{(K-2)(K-2(1-\eta^{-1}))},$$

which holds for any K if and only if $\zeta \leq 0$, which combined with the condition $\zeta \geq 0$, implies that $\zeta = 0$ is the only case where $\alpha \in [0,1)$ and $\frac{\partial \alpha}{\partial \eta} \geq 0$.

But notice that if $\zeta = 0$, then

$$\alpha = \frac{1-\eta^{-1}}{2} \leq \frac{1}{2}.$$

Thus, the Kimball demand also fails to generate a degree of strategic complementarity as high as the average of 0.9 in the data, while keeping the properties $\alpha \in [0,1)$ and $\frac{\partial \alpha}{\partial \eta} \geq 0$.

B.4 Proofs of Propositions for the Dynamic Model

Proof of Proposition 4.

Recall from Equation (15) that

$$\alpha = \frac{\varepsilon_D^{\varepsilon}}{\varepsilon_D^{\varepsilon} + \varepsilon_D - 1},$$

where ε_D is a firm's elasticity of demand and $\varepsilon_D^{\varepsilon}$ is its super-elasticity of demand in a symmetric point. Given the form of elasticities

$$\varepsilon_D(P_{j,k,t}, P_{j,-k,t}) = \eta - (\eta - 1)K^{\xi} \left(\frac{P_{j,k,t}^{1-\eta}}{\sum_{k \in K} P_{j,k,t}^{1-\eta}} \right)^{1+\xi},$$

we have $\varepsilon_D = \eta - K^{-1}(\eta - 1)$. Moreover,

$$\varepsilon_D^{\varepsilon} = \left[\frac{\partial \varepsilon_D(P_{j,k,t}, P_{j,-k,t})}{\partial P_{j,k,t}} \frac{P_{j,k,t}}{\varepsilon_D(P_{j,k,t}, P_{j,-k,t})} \right]_{P_{j,k,t}=1, \forall k} = \frac{(\eta-1)^2(1+\xi)(K-1)K^{-2}}{\eta - K^{-1}(\eta-1)}.$$

Plug these into the derivation for α and we get

$$\alpha = \frac{(1+\xi)(1-\eta^{-1})}{K + \xi(1-\eta^{-1})}.$$

Q.E.D.

Proof of Proposition 5.

This proof is an adaptation of the result in Lemma (8) 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,

$$S_{j,k}^{t-1} = (\ldots, S_{j,k,t-3}, S_{j,k,t-2}, S_{j,k,t-1}),$$

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:

$$\varsigma_t = (S_{j,k,t} \subset \mathcal{S}^t, p_{j,k,t} : S_{j,k,t}^t \to \mathbb{R})_{(j,k)\in J\times K},$$

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 $p_{j,k,t}(S_{j,k}^t)$ is linear in the vector $S_{j,k}^t$. This result follows with an argument similar to Lemma (3). Given this, let $p_{j,k,t}(S_{j,k}^t) = \sum_{\tau=0}^{\infty} \delta_{j,k,t}^{\tau} S_{j,k,t-\tau}$ 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 $\mathcal{S}^t$, $p_{j,k,t}(S_{j,k}^t) \in \mathcal{S}^t$, $\forall (j,k) \in J\times K$.

Now, pick a particular firm j,k and let $\varsigma_{-(j,k),t}$ denote the signals and pricing strategies that ς_t implies for all other firms in the economy except for j,k. Similar to Subsection A.4 let

$$\theta_{j,k,t}(\varsigma_{-(j,k),t}) \equiv (q, (p_{j,l,t}(S_{j,l}^t))_{l\neq k}, (p_{m,n,t}(S_{m,n}^t))_{m\neq j, n\in K})'$$

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

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

Since $\beta = 0$, firm j,k's problem is

$$\min_{S_{j,k,t}\subset\mathcal{S}^t} var(\mathbf{w}'\theta_{j,k,t}(\varsigma_{-(j,k),t})|S_{j,k}^t)$$
$$s.t. \quad \mathcal{I}(S_{j,k,t}, \theta_{j,k,t}(\varsigma_{-(j,k),t})|S_{j,k}^{t-1}) \leq \kappa.$$

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

$$p_{j,k,t}(S_{j,k}^t) = \mathbf{w}'\mathbb{E}[\theta_{j,k,t}(\varsigma_{-(j,k),t})|S_{j,k}^t].$$

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 $\theta_{j,k,t}(\varsigma_{-(j,k),t})$ and $S^t_{j,k}$ are Gaussian, $p_{j,k,t}(S^t_{j,k}) = \sum \delta^t_{j,k,\tau} S_{j,k,t-\tau}$ by Kalman filtering. Here for any $S_{j,k,t-\tau}$ that is not a singleton, let $\delta^t_{j,k,\tau}$ be a vector of the appropriate size that is implied by Kalman filtering. Therefore, by definition of $\mathcal{S}^t$, $p_{j,k,t}(S^t_{j,k}) \in \mathcal{S}^t$, meaning that there is a signal in $\mathcal{S}^t$ that directly tells firm j,k what their price would be under $S^t_{j,k}$ and $\varsigma_{-(j,k),t}$. Let $\hat{S}^t_{j,k} \equiv (S^{t-1}_{j,k}, p_{j,k}(S^t_{j,k}))$ and observe that by definition of $p_{j,k,t}(S^t_{j,k})$

$$var(\mathbf{w}'\theta_{j,k,t}(\varsigma_{-(j,k),t})|S^t_{j,k}) = var(\mathbf{w}'\theta_{j,k,t}(\varsigma_{-(j,k),t})|\hat{S}^t_{j,k}).$$

Therefore, we have found a single signal that implies the same loss for firm j,k under $S^t_{j,k}$. Now, we just need to show that it is feasible, which is straight forward from data processing inequality: since $p_{j,k,t}(S^t_{j,k})$ is a function $S^t_{j,k}$, we have

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

which concludes the proof for sufficiency of one signal.

Now, given $S^{t-1}_{j,k}$ and $\theta_{j,k,t}(\varsigma_{-(j,k),t})$ let $\Sigma_{j,k,t|t-1} \equiv var(\theta_{j,k,t}(\varsigma_{-(j,k),t})|S^{t-1}_{j,k})$. Without loss of generality assume $\Sigma_{j,k,t|t-1}$ is invertible. If not, then there are elements in $\theta_{j,k,t}(\varsigma_{-(j,k),t})$ that are colinear conditional on $S^{t-1}_{j,k}$, in which case knowing about one completely reveal the other; this means we can reduce $\theta_{j,k,t}(\varsigma_{-(j,k),t})$ 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

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

where $\mathbf{z}_t \equiv \frac{cov(S_{j,k,t}, \theta_{j,k,t}(\varsigma_{-(j,k),t})|S^{t-1}_{j,k})}{\sqrt{var(S_{j,k,t}|S^{t-1}_{j,k})}}$.The capacity constraint of the agent becomes

$$\mathbf{z}'_t \Sigma^{-1}_{j,k,t|t-1}\mathbf{z}_t \leq \lambda \equiv 1 - 2^{-2\kappa}.$$

Moreover, notice that the loss of the firm becomes

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

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 covarinace. 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

$$\max_{\mathbf{z}_t}(\mathbf{w}'\mathbf{z}_t)^2$$
$$s.t. \quad \mathbf{z}_t'\Sigma_{j,k,t|t-1}^{-1}\mathbf{z}_t \leq \lambda.$$

By Cauchy-Schwarz inequality we know

$$\begin{aligned}(\mathbf{w}'\mathbf{z}_t)^2 &\leq (\mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w})(\mathbf{z}_t'\Sigma_{j,k,t|t-1}^{-1}\mathbf{z}_t)\\ &\leq \lambda\mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w},\end{aligned}$$

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-Shwarz 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 problem.[50]

Now, I just need to show that a signal exists in $\mathcal{S}^t$ that implies this $\mathbf{z}_t^*$. To see this let

$$S_{j,k,t}^* = \mathbf{w}'\theta_{j,k,t}(\varsigma_{-(j,k),t}) + e_{j,k,t},\, e_{j,k,t} \perp (\theta_{j,k,t}(\varsigma_{-(j,k),t}), S_{j,k}^{t-1}),\, var(e_{j,k,t}) = (\lambda^{-1}-1)\mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w}.$$

Notice that $S_{j,k,t}^* \in \mathcal{S}^t$, and that

$$\frac{cov(S_{j,k,t}^*, \theta_{j,k,t}(\varsigma_{-(j,k),t})|S_{j,k}^{t-1})}{\sqrt{var(S_{j,k,t}^*|S_{j,k}^{t-1})}} = \sqrt{\frac{\lambda}{\mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w}}}\Sigma_{j,k,t|t-1}\mathbf{w},$$

meaning that $S_{j,k,t}^*$ solves the firm's problem. The optimal pricing strategy of the firm is

$$p_{j,k,t}^*(S_{j,k}^t) = \mathbb{E}[\mathbf{w}'\theta_{j,k,t}(\varsigma_{-(j,k),t})|S_{j,k}^t].$$

Finally, to get the form of the signal as shown in the Proposition simply plug in the vectors $\mathbf{w}$ and $\theta_{j,k,t}(\varsigma_{-(j,k),t})$ to get

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

[50]This solution can also be obtained by applying the Kuhn-Tucker conditions.

Q.E.D.

Proof of Proposition 6.

From the proof of Proposition 5 recall that in the equilibrium, for all $(j,k) \in J \times K$,

$$p_{j,k,t}(S^t_{j,k}) = \mathbf{w}'\mathbb{E}[\theta_{j,k,t}(\varsigma_{-(j,k),t})|S^t_{j,k}]$$

where $S^t_{j,k} = (S^{t-1}_{j,k}, S_{j,k,t})$ and $S_{j,k,t} = (1-\alpha)q_t + \alpha\frac{1}{K-1}\sum_{l\neq k} p_{j,l,t}(S^t_{j,l}) + e_{j,k,t}$. Now, from Kalman filtering

$$\mathbf{w}'\mathbb{E}[\theta_{j,k,t}(\varsigma_{-(j,k),t})|S^t_{j,k}] = \mathbb{E}[\mathbf{w}'\theta_{j,k,t}(\varsigma_{-(j,k),t})|S^{t-1}_{j,k}] + \frac{\mathbf{w}'cov(S_{j,k,t},\theta_{j,k,t}(\varsigma_{-(j,k),t}))}{var(S_{j,k,t}|S^{t-1}_{j,k})}(S_{j,k,t} - \mathbb{E}[S_{j,k,t}|S^{t-1}_{j,k}]).$$

Notice from the proof of Proposition 5 that

$$\frac{\mathbf{w}'cov(S_{j,k,t},\theta_{j,k,t}(\varsigma_{-(j,k),t}))}{var(S_{j,k,t}|S^{t-1}_{j,k})} = \frac{\lambda}{\mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w}}\mathbf{w}'\Sigma_{j,k,t|t-1}\mathbf{w} = \lambda.$$

Thus, using $p_{j,k,t}$ as shorthand for $p_{j,k,t}(S^t_{j,k})$,

$$p_{j,k,t} = (1-\lambda)\mathbb{E}[S_{j,k,t}|S^{t-1}_{j,k}] + \lambda S_{j,k,t}.$$

Finally, notice that $p_{j,k,t-1} = \mathbb{E}[S_{j,k,t-1}|S^{t-1}_{j,k}]$. Subtract this from both sides of the above equation to get

$$\pi_{j,k,t} \equiv p_{j,k,t} - p_{j,k,t-1} = (1-\lambda)\mathbb{E}[\Delta S_{j,k,t}|S^{t-1}_{j,k}] + \lambda(S_{j,k,t} - p_{j,k,t-1}),$$

where $\Delta S_{j,k,t} = S_{j,k,t} - S_{j,k,t-1}$. Now, subtract $\lambda\pi_{j,k,t}$ from both sides and divide by $(1-\lambda)$ to get

$$\pi_{j,k,t} = \mathbb{E}[\Delta S_{j,k,t}|S^{t-1}_{j,k}] + \frac{\lambda}{1-\lambda}(S_{j,k,t} - p_{j,k,t}).$$

Averaging this equation over all firms gives us the Phillips curve. To derive it, I take the average of every term separately and then sum them up.

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

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,

$$\frac{1}{JK}\sum_{(j,k)\in J\times K}(S_{j,k,t} - p_{j,k,t}) = (1-\alpha)q_t + \alpha\underbrace{\frac{1}{JK}\sum_{(j,k)\in J\times K}\frac{1}{K-1}\sum_{l\neq k}p_{j,l,t} - \frac{1}{JK}\sum_{(j,k)\in J\times K}p_{j,k,t}}_{=(\alpha-1)\frac{1}{JK}\sum_{(j,k)\in J\times K}p_{j,k,t}} + \underbrace{\frac{1}{JK}\sum_{(j,k)\in J\times K}e_{j,k,t}}_{\approx 0}.$$

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,

$$\frac{1}{JK}\sum_{(j,k)\in J\times K}(S_{j,k,t} - p_{j,k,t}) = (1-\alpha)(q_t - p_t) = (1-\alpha)y_t.$$

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

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

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

Q.E.D.

Proof of Proposition 7.

By the assumption of an impulse response, it corresponds to a sequence of signals that are zero until the initial time 0 when a shock hits. Therefore, at this initial time

$$\overline{\mathbb{E}^{j,k}_{-1}[\Delta q_t]} = \overline{\mathbb{E}^{j,k}_{t-1}[\pi_{j,-k,t}]} = 0.$$

Hence, the Phillips curve implies

$$\pi_0 = (1-\alpha)\frac{\lambda}{1-\lambda}y_0.$$

Moreover, $1 = \Delta q_0 = \pi_0 + \Delta y_0$. But $\Delta y_0 = y_0$ as $y_{-1} = 0$. Plug this into the above equation to get

$$\begin{aligned} \pi_0 &= (1-\alpha)\frac{\lambda}{1-\lambda}(1-\pi_0) \\ &= \frac{(1-\alpha)\lambda}{1-\alpha\lambda}. \end{aligned}$$

Additionally,

$$y_0 = 1 - \pi_0 = \frac{1-\lambda}{1-\alpha\lambda}.$$

Q.E.D.

B.5 The Symmetric Stationary Equilibrium and the 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

$$\Delta q_t = \sum_{n=0}^{\infty} \psi_q^n u_{t-n}.$$

Now, following Proposition 5 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

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

To characterize the equilibrium, 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, $\sum_{n=0}^{\infty} b^n_{j,k,t} v_{j,k,t-n}$ 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 $(S^{-1}_{j,k})_{(j,k)\in J\times K}$, suppose a strategy profile $(S_{j,k,t} \in \mathcal{S}^t, p_{j,k,t} : S^t_{j,k} \to \mathbb{R})_{k\in K, t\geq 0}$ 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

$$\begin{aligned} S_{j,k,t} &= (1-\alpha)q_t + \alpha \frac{1}{K-1} \sum_{l\neq k} p_{j,k}(S^t_{j,k}) + e_{j,k,t} \\ &= \sum_{n=0}^{\infty} \left[((1-\alpha)\psi^n_q + \alpha a^n)\tilde{u}_{t-n} + \alpha b^n \frac{1}{K-1} \sum_{l\neq k} v_{j,l,t-n} + e_{j,k,t} \right], \end{aligned}$$

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 a

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

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 $e_{j,k,t-n} \perp v_{j,l,t-n}, \forall l \neq k$. 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 let for a representative firm j, k

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

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. $\forall n \in \{0, \ldots, \bar{T}-1\}$, 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 $(a^n, b^n)_{n=0}^{\bar{T}-1}$.