Optimal Prices in Multisector Models under Rational Inattention

Chad Fulton

October 27, 2015*

Abstract

Recently empirical research has emphasized that theoretical models of price-setting must distinguish between the effects of aggregate and sector-level shocks, and moreover that they must support heterogeneous behavior across sectors. This paper develops a model that can deliver these features by extending the rational inattention price-setting approach pioneered by Maćkowiak and Wiederholt (2009) to a multisector setting. Our analytic solution to a special case of the rational inattention problem allows us to detail attention allocation mechanisms and explore implications. More generally, we find that the multisector setting preserves the desirable characteristic that firms respond differently to different types of shocks, allows for heterogeneous responses, and reduces the need for extreme calibration of key parameters.

*The most current version of this paper is available at http://pages.uoregon.edu/cfulton/research.html

1 Introduction

A large branch of macroeconomic literature is concerned with the apparent non-neutrality of monetary policy in the short-run, addressing the question of why *nominal* changes have *real* effects.[1] This literature stretches back to Keynes who suggested wage and price stickiness as a mechanism by which an economy might fail to fully and immediately adjust to nominal changes - and, thus, why an economy might operate out-of-equilibrium. Speaking somewhat loosely, if not all prices and wages adjust each period - whether due to existing contracts, menu costs, informational costs, etc. - then, for example, it may be in a recession that wages are "stuck" too high relative to their "natural" level resulting in unemployment "stuck" too high until such time passes that wages adjust, at which point the economy returns to long-run equilibrium.

This paper augments a typical general equilibrium model with monopolistically competitive intermediate goods firms by considering a multisector extension in which prices are fully flexible but firms face uncertainty about aggregate variables and real marginal costs. Following Maćkowiak and Wiederholt (2009), this uncertainty is modeled with rationally inattentive firms and results in a delay in the response of prices to monetary policy shocks. The baseline multisector model and an extension with relative demand shocks and intermediate production inputs provide additional targets for the attention of firms. Compared to Maćkowiak and Wiederholt (2009), we are able to relax the need for extreme calibration of volatilities in order to achieve a reasonable delay.

Contemporary sticky price general equilibrium models often introduce such stickiness by assuming a la Calvo (1983) that monopolistically competitive firms may only adjust their price each period with some constant exogenous probability related to the length of time since the adjustment (more generally, in these so-called "time-dependent" models the probability of adjustment may be related to the length of time since the previous adjustment). Other mechanisms include "state-dependent"

[1] This question can equivalently be put in terms of aggregate demand (why can governments manipulate aggregate demand in order to produce short-run economic effects?), aggregate supply (why is there an upward sloping short-run aggregate supply curve?), or the Phillips curve (why is there a short-run trade-off between inflation and output?).

models in which the probability that a firm may adjust their price depends also on the state of the economy (see for example Dotsey et al., 1999); the introduction of fixed costs associated with price-setting (see Golosov and Lucas Jr., 2007 for a recent example of this "menu-cost" approach); or the assumption that prices are fixed due to contracts (see for example Chari et al., 2000).

There exists another branch of the literature, however, that addresses the issue of short-run monetary non-neutralities by focusing instead on informational limitations of agents, so that the impediment to full adjustment is not a restriction on the flexibility to *perform* such changes, but rather a restriction on the information that would *prompt* change.[2] This idea was famously described by Phelps (1968) and Friedman (1968) and soon after was formalized by Lucas (1972) in the eponymous Lucas Islands model, in which agents face a signal-extraction problem to distinguish aggregate from idiosyncratic conditions. More recent incarnations of this idea can be found in Mankiw and Reis (2002), in which agents have a Calvo-like probability of receiving new information each period (the "sticky information" approach); Woodford (2001), in which agents face a dynamic signal-extraction problem; Angeletos and La'O (2010), in which the focus is on the *heterogeneity* of information imperfections across agents takes pride of place; and Maćkowiak and Wiederholt (2009), in which, following Sims (2003) in a a so-called "rational inattention" model, agents must split a limited amount of attention between observing aggregate and idiosyncratic conditions. It is within this literature that the current paper falls, following the rational inattention approach.

While many of the above mechanisms rely to at least some degree on an *ad hoc* imposition of sub-optimal behavior, in the seminal work on rational inattention Sims (2003) lays out a framework for microfoundations of imperfect information. While agents are still fully optimizing, in Sims' model they face an information-processing capacity constraint. Recognizing that they cannot pay attention to everything - that their information must necessarily be imperfect - they optimally use what capacity they do have.

We pursue this approach in this paper. In the perfect-information case firms set prices as a markup

[2]See Mankiw and Reis (2010) for a summary of the recent literature.

over nominal marginal costs, a step which requires knowledge of the contemporaneous aggregate price level, one's own contemporaneous productivity shock, and, due to the effect on aggregate demand, the contemporaneous shocks to every other firm. In the imperfect-information case, firms' uncertainty is formalized in terms of an information-processing capacity constraint, requiring firms to optimally divide their attention between aggregate and idiosyncratic shocks. In the special case of Gaussian white noise shocks, we derive analytic results on the optimal allocation of attention. More generally, the equilibrium behavior allows heterogeneous behavior related to firm- and sector-level characteristics.

This line of research is motivated by recent empirical work - in particular Bils and Klenow (2004), Golosov and Lucas Jr. (2007), Klenow and Kryvtsov (2008), and Boivin et al. (2009) - that suggests first that the results from traditional models of inducing monetary non-neutralities are not consistent with all observed inflation dynamics, and second that because disaggregated price series display markedly different inflation dynamics than do aggregated series, a successful model must begin work at the level of individual sectors. The former suggestion motivates the use of rational inattention as the key deviation allowing monetary non-neutralities - a suggestion which was also made in Sims' original work and has been already been followed up on, in particular in Maćkowiak and Wiederholt (2009). The latter implies that special attention must be paid to modeling sectors themselves; pursing this is one of the contributions of this paper.

Finally, this paper is especially related to three recent theoretical models. Woodford (2001), key in the recent revival of imperfect information models, describes a one-sector model in which firms face a dynamic signal-extraction problem. Whereas it *motivates* information imperfections by appealing to rational inattention, this paper (as does Maćkowiak and Wiederholt, 2009, below) *derives* the imperfections from optimizing behavior. Angeletos and La'O (2010) use a multisector real business cycle model to emphasize that heterogeneity of information can generate business cycles. The current paper's multisector approach and its focus on sectoral heterogeneity in particular follows from this realization. Finally, this paper can be thought of most directly as an extension of Maćkowiak and Wiederholt (2009) who present a one-sector rational inattention model and derive conditions for optimal allocation in the cases of Gaussian white-noise and stationary shocks. In the stationary case they describe the implications for inflation dynamics but are forced to use computational techniques. A one-sector special case of the current paper's primary result reduces to the model found in section IV of their work.

The remainder of this paper is structured as follows. In section 2, the literature is reviewed in some detail. Section 3 introduces information theory and details some useful results. Section 4 presents the model, section 5 the equilibrium conditions, and section 6 the results. Section 7 concludes.

2 Related Literature

With the threefold goal of (1) positioning the current paper along the arc of the existing literature, (2) explaining the relationship between this model and closely related models, and (3) describing relevant features of the data that inform modeling choices, this section proceeds by briefly describing the imperfect information literature, presenting related empirical research, and introducing recent theoretical models against which this paper's model will be contrasted.

2.1 Lucas (1972, 1973)

Robert Lucas, Jr. laid out the first formal models with imperfect information leading to monetary non-neutrality using geographically separated islands as the device preventing perfect information. Agents on each island receive signals about unknown variables and must solve a (static) signal extraction problem each period to distinguish idiosyncratic from aggregate conditions, with the key result that individuals' misperceptions of movements in nominal price for movements in real price can allow monetary policy to affect real variables. In particular, if firms (individuals in the model) mistake a purely nominal increase in the price level for an increase in their real price they will increase employment. While these simple models have been superseded, the insight that individuals' optimizing behavior depends on aggregate variables that may be unknown permeates all of the subsequent imperfect information literature, and will be a central focus in the model presented in the current paper.

Despite a high level of subsequent interest in imperfect information models, the literature largely died out by the early 1990s due especially to several critiques that could not at the time be overcome. The first was the difficulty of squaring the model with data - for example, the model required all information to become available after one period, implying that effects of monetary policy would not be more persistent than that delay. However, "periods" long enough to match the observed persistence of monetary policy effects implied an implausibly long delay before agents were made aware of that policy. Second, technical difficulties arose in modeling higher order expectations (see Townsend, 1983).

Strategic effects between individuals, a feature not highlighted in Lucas' models, have since become important and are considered in more recent imperfect information models as well as in models with other mechanisms inducing stickiness; for example, the implications of pricing decisions as strategic substitutes or as strategic complements are important to New Keynesian models (see Woodford, 2003 sections 3.1.3 - 3.1.4). Strategic considerations are similarly important in Angeletos and La'O (2010), described below, and will aid interpreting the current paper's results.

2.2 Morris and Shin (2002)

Stephen Morris and Hyun Song Shin (2002) were instrumental in restarting the discussion of imperfect information models, demonstrating welfare implications of imperfect and heterogeneous information in the presence of strategic complementarities and drawing out the link between higher order expectations and strategic behavior. In their model, agents receive public and private signals about unknown variables; one important result is that increasing precision of public signals may actually reduce welfare. While they do not focus specifically on pricing decisions, they show that Lucas' model is equivalent to the one they consider. Their intuition and solution techniques carry over to a wider range of modeling approaches, including the multisector model in Angeletos and La'O (2010) and the current paper.

2.3 Sims (2003)

Chris Sims (2003) introduced rational inattention, a modeling paradigm in which rational optimizing agents could fail to take into account even freely available information, thus providing microfoundations for information imperfections.[3] His suggestion provides a response to one critique of the Lucas model: if agents do not pay attention to monetary policy, it does not matter how quickly the information is made available. The technical component of these models introduces information theory to economics, a topic which is described below in some detail (see *Information Theory*).

Notice that many of the papers described here use "signals" received by agents as the technical device encoding imperfect information. Sims shows that in a special case - a linear quadratic optimization problem along with Gaussian stochastic processes - rational inattention can lead to results that are identical to simply assuming agents receive noisy signals, but the rational inattention approach provides a framework for the optimal selection of signals by agents and shows how the noise varies systematically with underlying model parameters.

The central contribution of the current paper is the solution of a rational inattention problem by intermediate goods firms in the presence of both idiosyncratic and aggregate shocks.

2.4 Woodford (2001)

A paper prepared by Michael Woodford for a conference commemorating Edmund Phelps was similarly important in the revival of the imperfect information literature. It extended Lucas' model to one in which information does not become available after a one period delay, so that individuals

[3] For more on modeling rational inattention, see Sims (1998), Sims (2005), and Sims (2010).

face a dynamic signal extraction problem. Agents are also less informed than in Lucas' model: not only are they unaware of aggregate conditions, they are also unaware of other agents' expectations of aggregate conditions (and their expectations of expectations of ...). Woodford shows in a one-sector model that results hinge on an infinite sum of higher order expectations and uses the Kalman filter to solve the (dynamic) signal extraction problem. The key result, driven largely by sluggishness in the response of higher-order expectations, is that the real effects of monetary policy can persist for an arbitrary number of periods.

In solving the model, the assumption is made that individuals are given signals about aggregate quantities. To justify it, Woodford briefly refers to Sims (2003) but does not explore the rational inattention problem. Pursuit of microfoundations for the optimal selection of signals by agents has been an area of subsequent research; one example is Maćkowiak and Wiederholt (2009) who solve the rational inattention optimal price problem for a one-sector model comparable to Woodford's. Their paper is described in detail below, and the current paper is a partial extension of their results to a multisector setting.

2.5 Angeletos and La'O (2010)

While imperfect information has traditionally been an amplification mechanism for monetary policy, Angeletos and La'O (2010) consider its ability to induce business cycles in a purely real setting. They develop a multisector model in which islands provide boundaries to information dispersion and in which intermediate goods firms set quantities (rather than prices, as has been the case above). They show that it is the heterogeneity of information across the islands rather than the magnitude of imperfection that drives their results.

They find that dispersed (and heterogeneous) information can lead to fluctuations and inertia in macroeconomic variables and that the generated fluctuations match qualitative facts about business cycles that other imperfect information models cannot (although they do not pursue any quantitative investigation). Furthermore, they emphasize that the model can generate these fluctuations even when individuals are nearly perfectly informed, so long as information is dispersed.

Strategic complementarities (although not those of the New Keynesian type, as they point out) induced by "trade linkages" (the between-island elasticity of substitution) are important in understanding the interdependence of firms' decisions and are crucial to their results. In particular, it is only when firms' decisions are complementary (goods across islands are not perfect substitutes) that imperfect information has real effects.

Their paper provides both the theoretical motivation and basic model setup for the current paper, although here we return to the consideration of price-setting firms. Their emphasis on the importance of modeling the interactions of heterogeneous agents in the presence of imperfect information provides an impetus for the extensions of Maćkowiak and Wiederholt (2009) that this paper considers (this point discussed at greater length below).

2.6 Mackowiak and Wiederholt (2009)

Bartosz Mackowiak and Mirko Wiederholt (2009) (hereafter MW) consider optimal price-setting behavior of rationally inattentive firms in the face of idiosyncratic and aggregate shocks, finding that for certain calibrations (in which idiosyncratic shocks are an order of magnitude more volatile than aggregate shocks) the real effects of monetary policy can persist for an arbitrary number of periods, a result that hinges on agents choosing not to pay much attention to monetary policy.

They first consider a special case of the model in which stochastic processes are Gaussian white noise and in which case an analytical result may be found. Although this case does not induce persistence in the model (since all shocks are purely transitory), it draws out the intuition of rational inattention and shows how the firms' attention allocation decision depends on model parameters. One interesting point involves strategic complementarities: strategic complementarities in price setting imply strategic complementarities in information acquisition. It is this portion of their paper which the current paper extends to a multisector setting.

Second, they consider more realistic stochastic processes for shocks, which sufficiently complicates results that computational techniques must be used to solve the information allocation problem, and show that their model can generate real effects of monetary policy and explain why firms might respond quickly to idiosyncratic shocks but slowly to aggregate shocks. In order to do this, their model requires calibrating the volatility of idiosyncratic shocks to be at least an order of magnitude larger than the volatility of aggregate shocks. While the data suggests that idiosyncratic shocks are more volatile than aggregate shocks, it does not support this degree of difference (see the discussion of the related empirical work, below, for details). One of the contributions of this paper is relaxing the differential in volatility required to achieve the appropriate behavior.

A more complete discussion of the contributions of the current paper *vis a vis* MW follows a brief summary of Maćkowiak et al. (2009), below.

2.7 Boivin et al. (2009)

Jean Boivin et al. (2009) use a factor augmented vector autoregression (FAVAR) approach to estimate separately the effects of aggregate and idiosyncratic disturbances on price-setting behavior, finding that prices are flexible with respect to idiosyncratic shocks but sticky with respect to aggregate (in particular monetary) shocks, a feature of the data which they suggest is not consistent with many contemporary stickiness-inducing mechanisms (for example they suggest that the Calvo mechanism cannot explain the flexibility with respect to idiosyncratic disturbances). They lay out seven stylized facts, all of which provide a rich research agenda with respect to the theoretical modeling of imperfect information, and two of which may justify modeling decisions in this paper.

First, their primary result is the importance of distinguishing between idiosyncratic and aggregate components of price change; this is of central consideration in the rational inattention approach to modeling optimal pricing decisions. Note that this result is not specific to Boivin et al. (2009) but is also borne out in related empirical efforts.[4] Second, they suggest that idiosyncratic shocks

[4]See in particular Bils and Klenow (2004) and Klenow and Kryvtsov (2008) for supporting empirical evidence.

driving price changes are supply shocks at the sectoral level, providing support for the approach in the current paper for integrating idiosyncratic shocks (as opposed to introducing them as, for example, demand shocks).

Their other facts have implications for extensions of the current paper and will be of particular importance when introducing persistence in shocks. Several are suggestive of interesting extensions (for example, including a further integration of the effects of market power, see below), while several others appear to present challenges to the current approach. For example, they suggest that the reaction to all types of shocks is faster in sectors with more volatile idiosyncratic shocks, whereas the results of MW specifically suggest that increased volatility in idiosyncratic shocks reduces the reaction to aggregate shocks.

Finally, the FAVAR approach allows them to estimate the relative volatilities of the aggregate and idiosyncratic components. They find that "while the mean volatility of the common component of inflation lies at 0.33 percent, the volatility of the sector-specific component is more than three times as large". While this differential is not enough to generate the appropriate behavior in the single-sector model of MW, it can be enough in the multisector model presented here. Thus our model insulates the rational inattention approach from criticism that extreme calibrations are required.

2.8 Mackowiak et al. (2009)

Mackowiak et al. (2009) (hereafter MMW) consider a simple multisector extension to their model in MW as part of an effort to compare the ability of several mechanisms for stickiness to match empirical results. The mechanisms they consider including the Calvo model, sticky information, menu costs, and rational inattention. The stylized facts they attempt to match are in the same vein as those described by Boivin et al. (2009), and they find that the rational inattention model is best able to fit them. There are several key points that distinguish the approach of MW and MMW from that of the current paper.

In both papers, Mackowiak et al. consider the tradeoff firms face between paying attention to aggregate conditions or idiosyncratic conditions, where idiosyncratic conditions refers specifically to the firm's own productivity shock. Productivity shocks to other firms either net to zero in aggregate, as in MW, or are simply grouped with other aggregate shocks (for example monetary policy incorporated via aggregate demand shocks), as in MWW. This leads to simpler optimization problems as firms only spread their attention between two signals and still distinguishes nominal (monetary policy) from real (productivity) shocks.

In contrast, the current paper considers firms' attention problems as between all shocks separately; this is desirable for several reasons. First, it embraces the emphasis in Angeletos and La'O (2010) of the importance of considering how interactions of heterogeneous firms with dispersed information alone can generate real effects; they note that "even if one is ultimately interested in a monetary model, understanding the positive and normative properties of its underlying real backbone is an essential first step".

Second, this paper is meant to provide a baseline model for considering more complicated models of firm interactions and allocation problems, and for introducing firm-level heterogeneity in pricing decisions. For example, it would make sense that firms are more easily able to pay attention to shocks of more closely related sectors. Integrating this requires a model like that of the current paper. Another motivating example can be found in the stylized facts of Boivin et al. (2009), who find that the speed of reaction to monetary policy shocks is related to the degree of monopoly power enjoyed by firms, an extension which will require attention to the underlying heterogeneity of firms' attention allocation problems.

2.9 De Graeve and Walentin (2014)

Two stylized facts derived in both Boivin et al. (2009) and Maćkowiak et al. (2009) are that aggregate shocks display substantial persistence but are characterized by low volatility, and that idiosyncratic shocks have low persistence but high volatility. It is partially these facts that validate the rational inattention approach. More recently, De Graeve and Walentin (2014) suggests that if the effects of measurement error are properly accounted for, these stylized facts may be overturned. In particular, they suggest that idiosyncratic shocks may be persistent and have low volatility, potentially presenting a problem for the rational inattention approach. The multisector model presented here allows the rational inattention approach to respond.

First, the authors note that their estimation approach does not allow them to distinguish between measurement error and additional structural shocks, but they argue that the rational inattention model in MW can not obviously account for the required additional shocks. By contrast, the multisector model developed below can easily accomodate the required additional shocks. Chapter 2 of this prospectus presents evidence that supports the features they identify as measurement error are more likely structural shocks affecting certain sectors.

Second, as discussed above, the multisector model reduces the volatility differential required by the rational inattention approach. Thus even if measurement error is distorting the stylized facts as they argue, rational inattention may still be used to explain observed price dynamics.

3 Information Theory

Rational inattention borrows from the theory of information and communication a mathematical model of information processing, and applies it to economic agents.[5] A telegraph wire serves as a channel through which a message passes from source as input to recipient as output. That it can only transmit a finite message in any given time interval is described as a finite Shannon "capacity".[6] In our model, an agent serves as a channel through which observations about the economy are translated into economic actions; the inability of the agent to process all information is modeled in terms of a finite Shannon capacity.

[5]See Sims (2003) or Sims (2010) for an introduction to rational inattention in economics, or Cover and Thomas (2006) for a general book length treatment of information theory. This section is largely drawn from chapters 2 and 9 of that work.

[6]The seminal work in information theory is Shannon (1948).

The basic quantity in information theory is entropy, a measure of the uncertainty associated with a random variable. Letting X denote a random variable with probability mass function or density P, entropy is defined as

$$H(X) = -E[\log(P(X))] \qquad \text{Entropy}$$

Notice that entropy is defined over probabilities and therefore must be positive and that is zero exactly when the distribution of X is degenerate. The units in which entropy is expressed depend on the base of the logarithm; typically "bits" are used, corresponding to log base 2. Two closely related quantities are joint entropy and conditional entropy which measure, respectively, the uncertainty of two random variables together and the uncertainty of a random variable conditional on the observation of another random variable. Letting S denote a second random variable, these quantities and their connection (called the "chain rule") are defined

$$H(X, S) = -E[\log(P(X, S))] \qquad \text{Joint entropy}$$

$$H(X|S) = -E[\log(P(X|S))] \qquad \text{Conditional entropy}$$

$$H(X, S) = H(X) + H(S|X) \qquad \text{Chain rule}$$

In the case that the two random variables are independent, the conditional entropy is identical to the (unconditional) entropy and so the joint entropy is the sum of the individual entropies.

Using these two definitions, we can define a quantity which will be of central interest in rational inattention, mutual information. The mutual information between two random variables X and S is the *reduction in uncertainty* about X given the observation of S; in that way it measures the information content contained in one variable about another. It is defined as

$$\mathcal{I}(X; S) = H(X) - H(X|S) \qquad \text{Mutual information}$$

In the case that the variables are independent so that is no reduction in uncertainty, then $I(X; S) =$

$H(X) - H(X) = 0$. Supposing that $\mathbf{X}$ and $\mathbf{S}$ are finite n-dimensional independent vectors such that X_i and S_j are independent if and only if $i \neq j$, then

$$\mathcal{I}(\mathbf{X}; \mathbf{S}) = \sum_{i=1}^{n} I(X_i; S_i)$$

This result is important because in our model we will consider attention allocation problems in which X and S will be vectors and our assumptions will make them internally independent though mutually dependent. Typically we will think of X as fundamentals of interest (for example a stochastic shock) and S as signals received by agents that provides some information about those fundamentals. In the rational inattention framework, agents optimally choose the signals, but must do so subject to an information constraint. That constraint is formalized as a maximum level of mutual information between the fundamental and the signal: $\mathcal{I}(X; S) \leq \kappa$. Since the fundamental is a vector, in addition to respecting the overall information-capacity constraint the agents must tradeoff between paying close attention to one variable or to another.

Notice that entropy and mutual information are scalar valued regardless of the dimension of the random variables, so that all of the data regarding uncertainty and information is expressed in a single number. It is this property that leads to the simplicity of the capacity constraint, which can be introduced into the model with only a single new free parameter, κ.

Despite the ease of modeling the constraint, calibrating the value of κ is difficult for two reasons. First, it is measured in bits, where 1 bit is the level of uncertainty related to a fair coin toss. This is a difficult to interpret quantity in the context of real-world economic decisions, and is complicated further by the simplification inherent to an economic model. Second, even supposing that the bit-value of actual information-processing capacity could be assessed, the model only captures one aspect of the decision problems facing a firm, so it is unclear what proportion of their attention is specifically devoted to the stochastic elements represented in the model.

Sims (2010) suggests that when a price is assigned to additional information capacity so that the amount is variable, agents choose a relatively small amount. In practice, these models are often calibrated such that agents set prices that are close to the optimum.

4 Model

There are a continuum of identical households $h \in H$ with associated measure μ_H, each of which consumes a continuum of differentiated goods $j \in J$ with associated measure μ_J. Consumers have nested constant elasticity of substitution (CES) preferences that induce a partition on the goods (alternatively firms) J into sectors $\{J_1, \ldots, J_I\}$, not necessarily of equal size. For convenience and so that aggregates can be identified with averages assume the total measure of households and goods is unity, so that $\mu_H(H) = \mu_J(J) \equiv 1$. To ease the remaining notation define $\mu_i \equiv \mu_J(J_i)$ as the size (measure) of sector i in the space of all firms.[7] Sectors will be typically indexed by i; l will also be used when multiple sector-level indices are required.[8]

In each period households consume, supply labor to each firm j[9], buy bonds, and receive profits based on their ownership in firms. For simplicity, each household owns an equal share of every firm. Firms set their prices each period to maximize the expected value to households of profits.

4.1 Households

Household utility is a discounted stream of expected utility, additively separable in time

$$\mathcal{U}\left(\{C_{hjt}, n_{hjt}\}_{j\in J, t\geq 0}\right) = E_0 \sum_{t=0}^{\infty} \beta^t \left[U\left(\{C_{hjt}\}_{j\in J}\right) - \int_J v(n_{hjt}) dj \right]$$

[7] In general a subscript j will refer to a firm $j \in J$, whereas the subscript i will denote a sector $J_i \subseteq J$. Since the sectors partition the set of goods, it is implicit that there is exactly one sector i corresponding to each firm j.

[8] The convention is that when j and i appear in the same equation, i refers to the industry such that $j \in J_i$. When a sector-level variable *not* referring to the sector of firm j is present, it will be indexed by l.

[9] This is merely for convenience. An equivalent setup has each household specializing in only one type of labor, see for example Woodford (2003) section 3.1.

The nested constant elasticity of supply (CES) preferences yield two Dixit-Stiglitz aggregators. The first level of aggregation describes "sector-level" goods[10]

$$C_{hit} = \left[\int_{J_i} \mu_i^{r-1} C_{hjt}^r dj \right]^{\frac{1}{r}} \tag{1}$$

where the associated within-sector elasticity of substitution is $\eta = \frac{1}{1-r}$. As usual, assume that goods are gross substitutes so that $\eta \in [1, \infty)$ and so $r \in [0, 1)$. The sector-level goods are then further aggregated into a "consumption good"

$$C_{ht} = \left[\sum_{i=1}^{I} \mu_i^{1-p} C_{hit}^p \right]^{\frac{1}{p}} \tag{2}$$

where the associated between-sector elasticity of substitution is $\rho = \frac{1}{1-p}$. Again assuming that goods are gross substitutes yields $\rho \in [1, \infty)$ and $p \in [0, 1)$. η and ρ are not constrained otherwise - for example they are not necessarily the same. Using this same basic setup, Angeletos and La'O (2010) describe the within-sector elasticity as the degree of market power of intermediate goods and the between-sector elasticity as a measure of trade linkages and strategic complementarities.

From this it is clear that the sizes of the sectors μ_i are defined entirely by consumers' relative demand weights. However, as in Woodford (2003) section 3.2.5, these weights can be reinterpreted as the product of a structural sector size parameter together with a relative demand weight.

Then given these aggregates, the households' optimization problems can be written

$$\max_{\{C_{ht}\}_{t \geq 0} \{n_{hjt}\}_{j \in J, t \geq 0}} = E_0 \sum_{t=0}^{\infty} \beta^t \left[u(C_{ht}) - \int_J v(n_{hjt}) dj \right]$$

where u is the instantaneous utility of consumption defined in terms of the consumption good and $v(n_{hjt})$ is the instantaneous disutility of labor. These are assumed to have the usual constant

[10] Notice that in the integral there appear both the indices i and j. Thus the i refers to the unique sector such that $j \in J_i$.

relative risk aversion forms

$$u(C) = \frac{C^{1-\sigma}}{1-\sigma} \quad \text{and} \quad v(n) = \frac{n^{1+\varepsilon}}{1+\varepsilon}. \tag{3}$$

Each period, the households' choices must satisfy the budget constraint

$$P_t C_{ht} + B_{ht+1} \leq \int_J \theta_{hj} \pi_{jt} dj + \int_J W_{jt} n_{hjt} dj + R_t B_{ht} \tag{4}$$

where C_{ht} is the consumption good, P_t is a price index corresponding to the cost-minimizing way to purchase one unit of the consumption good, B_{t+1} is a risk-free bond purchased in period t that yields income in period $t+1$ subject to the gross nominal risk free rate of return R_{t+1}, θ_{hj} is the share of firm j owned by the household, π_{jt} denotes profits from firm j, W_{jt} is the wage paid by firm j, and n_{hjt} is the labor provided by household h to firm j.

4.2 Firms

All intermediate goods firms produce differentiated output using a constant returns to scale technology with labor (denoted n_{jt}) as the sole input and a sector-specific productivity shock[11]

$$Y_{jt} = \varphi_{it} n_{jt} \tag{5}$$

For the moment we will remain agnostic about the variables present in the firms' information sets at time t. The nature of the shocks is discussed below. Assuming competitive factor markets, a firm's period profit is

$$\pi_{jt} = P_{jt} Y_{jt} - W_{jt} n_{jt} \tag{6}$$

[11] The model be expanded to allow a composite productivity shock, with firm-specific, sector-specific, and aggregate components, relative demand shocks, and intermediate inputs. A model with these extensions is introduced in *Extension: relative demand shocks and intermediate inputs*.

Firms have a degree of market power controlled by the within-sector elasticity η, and face inverse demand curves derived from households' optimizing behavior. Thus they choose prices so as to maximize the value of their profits to the owning households whose marginal utility of wealth is $u'(C_t)$; the intertemporal problem at time t for firm j is

$$\max_{\{P_{jt+s}\}_{s=0}^{\infty}} E_{jt} \sum_{s=0}^{\infty} \left\{ u'(C_t) \left[\prod_{l=1}^{s} \frac{1}{R_{t+l-1}} \right] \left(P_{jt+s} - \frac{W_{jt+s}}{\varphi_{it+s}} \right) Y_{jt+s} \right\}.$$

Since this model forgoes sticky prices in favor of informational frictions, firms re-optimize each period and need only solve the following static problem in each period separately

$$\max_{P_{jt}} E_{jt} \left[u'(C_t) \left(P_{jt} - \frac{W_{jt}}{\varphi_{it}} \right) Y_{jt} \right] \tag{7}$$

4.3 Government

Following the literature on imperfect information (Lucas, 1972, Woodford (2001), Mankiw and Reis (2002), Mankiw and Reis (2002)) we appeal to a quantity theory of money to specify a exogenous stochastic process for aggregate demand, assumed to be the result of monetary policy implemented by some policy instrument. This process will be the only aggregate shock to the economy.

$$Q_t = P_t Y_t$$

As formulated above, fiscal policy is excluded from the model to maintain focus on the firms' attention allocation problem, although in a very similar model Angeletos and La'O (2010) find that government intervention is only useful insofar is it can mitigate the distortionary effects of market power to improve efficiency, a topic not under central consideration here.

4.4 Stochastic processes

There are two exogenous stochastic processes to be specified, that for nominal aggregate demand $\{Q_t\}_{t=0}^{\infty}$ and that for idiosyncratic productivity shocks $\left\{[\varphi_{lt}]_{l=1}^{I}\right\}_{t=0}^{\infty}$. Here we assume the simple case that all processes are distributed log-normal in such a way that their logs are Gaussian white noise. All processes are assumed to be mutually independent. Formally the shocks are described

$$
\begin{aligned}
q_t &\equiv \log Q_t & q_t &\overset{iid}{\sim} N(0, \sigma_q^2) & & \qquad (8)\\
\phi_{it} &\equiv \log \varphi_{it} & \phi_{it} &\overset{iid}{\sim} N(0, \sigma_{\phi_i}^2) & l &= 1, \ldots, I
\end{aligned}
$$

Insofar as it would be difficult to argue that independent Gaussian white noise shocks represent the true stochastic nature of the economy, this specification is only a precursor to a more complete analysis attempting to match actual macroeconomic dynamics. Unfortunately models with more realistic stochastic processes do not admit analytic results and the Gaussian white noise case provides an illuminating special case in which to consider the attention allocation trade-offs faced by firms.

One could easily accomodate non-zero mean processes by redefining the above variables as be deviations from the mean. Processes with a deterministic trend could be similarly incorporated.

For notational convenience, collect the stochastic processes into an ordered tuple

$$\Omega = \{\{q_t\}, \{\phi_{1t}\}, \cdots, \{\phi_{It}\}\} \tag{9}$$

indexed by ω.

4.5 Imperfect Information

The expectations operator in the above formulation of an intermediate good firm's problem suggests that at least some contemporaneous economic variables - in this case aggregate consumption, firm-specific nominal wages, the sector-specific aggregate shock, and aggregate output - are unknown to the firm at the time they must set the price of their differentiated good. This raises two questions: (1) why would a firm be unaware of these (or any) contemporaneous conditions, and (2) which contemporaneous variables are unknown to the firm. Both of these questions have been of considerable interest in the literature on imperfect information, as described above.

Here we take the position that insofar as agents must process any information they wish to use and have a limited ability to assimilate even widely available information, all variables are *a priori* unknown. Agents only observe variables at all to the extent that they specifically allocate attention to do so. This is formalized in terms of the rational inattention framework of Sims (2003) with the assumption that agents have a finite information processing capacity κ. The imperfections are thus inherent to the agents and not the information itself; in fact, this approach requires that all the relevant information exists and is freely available.[12]

4.6 Signals

The device through which agents will receive (incompletely processed) information takes the form of signals $s_{jt}^{(\omega)}$ where j is the firm receiving the signal and ω indicates one of the stochastic processes described above. While in principle the space of possible processes among which agents may select signals is unrestricted with respect to distribution, in practice the structure of this problem - in particular the Gaussian white noise exogenous processes and a log-quadratic approximation to the profit function taken below - implies that optimal signals will be Gaussian.[13] For this

[12]See Sims (2005) for a discussion of this and related topics regarding the assumptions implicit to rational attention models.

[13]See Sims (2003) or Maćkowiak and Wiederholt (2009) for a proof of this result.

reason, we follow MW section IV in hereafter restricting the space of possible signals to those that are of the form "true state plus white noise"[14]

$$
\begin{aligned}
&s_{jt}^{(q)} = \tilde{q}_t + \psi_{jt}^{(q)} && s_{jt}^{(q)} \sim N(\tilde{q}_t, \sigma_q^2 + \sigma_{\psi_j^{(q)}}^2) && \\
&s_{jt}^{(l)} = \tilde{\phi}_{lt} + \psi_{jt}^{(l)} && s_{jt}^{(l)} \sim N(\tilde{\phi}_{lt}, \sigma_{\phi_l}^2 + \sigma_{\psi_j^{(l)}}^2) && l = 1, \ldots, I
\end{aligned}
\tag{10}
$$

The signals are written in terms of the deviation-from-mean forms to maintain the possibility of non-zero-mean processes even though here, for example, $\tilde{q}_t = q_t - Eq_t = q_t$. Although a formal connection will be derived below, a firm's attention problem can be informally described as the optimal reduction (or more properly selection) of the noise in signals subject to a constraint on the maximum amount of noise reduction across all signals.[15]

5 Equilibrium

An equilibrium is a collection of processes for consumption, labor, wages, prices, and signals

$$
\{C_{hjt}, n_{hjt}, W_{hjt}, P_{jt}, s_{jt}^{(\omega)}\}_{h,j,\omega,t}
$$

such that markets clear, households maximize utility, and firms (1) set optimal prices given available information, and (2) direct their attention such that the signals they receive about unknown quantities of interest are optimal. Notice that due to constant elasticity of substitution preferences, household optimization will uniquely define the processes for the aggregate price level $\{P_t\}$ and real aggregate demand $\{Y_t\}$, see (12) and (11) respectively. The full conditions governing optimal behavior are derived below.

[14] The use of the index l indicates that each firm j receives a separate signal for the shock to each industry.

[15] This model falls under the special Gaussian-linear-quadratic case. Sims (2010) section 3.2 presents this and related intuition for these types of models.

5.1 Market Clearing

The three markets in this model - goods, assets, and labor - yield three market clearing conditions. Since we have assumed constant elasticity of substitution preferences, we need only specify market clearing in terms of the consumption good, $C_t = Y_t$. If this holds, then in equilibrium (in particular given optimal household behavior) the demand functions for intermediate and sector-level goods, detailed below, guarantee market clearing at those levels. Since this model admits a representative household, no bonds will be purchased or sold in equilibrium so that the asset market clearing condition is $B_t = 0$ for all periods t. Finally, the labor market equilibrium requires $n_{jt} = \int_H n_{hjt} dh$.

5.2 Optimal Household Behavior

Standard results for constant elasticity of substitution preferences give demand functions for disaggregated goods in terms of aggregated quantities.[16] Since all households are identical their optimal behavior will be identical and can be analyzed as derived from a optimizing representative household. For this reason, in all of what follows we drop the household subscript.

$$C_{jt} = \frac{1}{\mu_i}\left(\frac{P_{jt}}{P_{it}}\right)^{\frac{1}{r-1}} C_{it} \qquad\qquad C_{it} = \mu_i \left(\frac{P_{it}}{P_t}\right)^{\frac{1}{p-1}} C_t \tag{11}$$

Corresponding to these demand functions are price indices prescribing the (minimal) cost of obtaining one unit of an aggregated quantity

$$P_t = \left[\sum_{i=1}^{I} \mu_i P_{it}^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \qquad\qquad P_{it} = \left[\int_{J_i} \frac{1}{\mu_i} P_{jt}^{\frac{r}{r-1}} dj\right]^{\frac{r-1}{r}} \tag{12}$$

Given these demand functions, the household's intertemporal problem can be analyzed in terms only of the consumption good. As usual, the solution is described by an Euler equation and a static

[16]See *A-1: Constant Elasticity of Substitution Preferences* for details.

first-order condition[17]

$$u'(C_t) = \beta E_t \left[R_{t+1} \frac{P_t}{P_{t+1}} u'(C_{t+1}) \right] \tag{13}$$

$$v'(n_{jt}) = \frac{W_{jt}}{P_t} u'(C_t) \tag{14}$$

5.3 Optimal Price Setting

Optimal behavior on the part of the firm will be considered in two stages. First, no matter the signals they actually receive about the state of the economy, firms must optimally set their decision variable given that information. In the case of perfect information this is the standard profit maximization problem faced by a monopolist. In the case of imperfect information a log-quadratic approximation to the profit function yields the certainty equivalence result that the optimal imperfect information price is simply the expectation of the optimal perfect information price.

Second, firms must select the signals they receive. Here they achieve this by allocating their attention such that they minimize the expected loss in profits from setting a non-optimal price subject to a constraint on the maximum attention they can spread across all variables.

Perfect Information

As a baseline, consider firms with perfect information. In this case there is no attention allocation so that the firms' entire problem reduces to the standard profit maximizing problem faced by a monopolist

$$\max_{P_{jt}} u'(C_t) \left(P_{jt} - \frac{W_{jt}}{\varphi_{it}} \right) Y_{jt}$$

[17] See *B-1: Optimal Household Behavior* for details.

This yields the standard result that monopolists set price as a markup over nominal marginal costs

$$P_{jt}^{\diamond} = \frac{1}{r}\frac{W_{jt}}{\varphi_{it}}$$

which can be rewritten in terms of model fundamentals and in the form of proportional (log) deviation from the point at which all prices are the same as[18]

$$\tilde{p}_{jt}^{\diamond} = -\gamma\tilde{\phi}_{it} + \zeta\tilde{q}_t + (1-\zeta)\tilde{p}_t \tag{15}$$

where $\tilde{q}_t$ represents nominal aggregate demand, $\zeta \equiv \alpha(\sigma+\varepsilon)$ relates to strategic complementarities between firms' pricing decisions, $\gamma \equiv \alpha(1+\varepsilon)$, and $\alpha \equiv (1+\rho\varepsilon)^{-1}$.[19] Integrating across all firms and applying a log-linear approximation yields the perfect-information equilibrium aggregate price

$$\tilde{p}_t^{\diamond} = \tilde{q}_t - \frac{\gamma}{\zeta}\sum_{l=1}^{I}\mu_l\tilde{\phi}_{lt} \tag{16}$$

Strategic complementarities

The firm's pricing rule (15) exposes $(1-\zeta)$ as a parameter governing strategic complementarities in the model. If it is positive (so $\zeta < 1$), firms' responses to changes in the aggregate price level will be complementary, whereas if it is negative, the aggregate price will act as a strategic substitute. This parameter appears in and has been important to both models with price stickiness and models with informational frictions. Typical calibrations put the value of ζ between 0.12 and 0.4, which implies that prices are strategic complements, and the parameter governing strategic complementarities, $(1-\zeta)$, is between 0.6 and 0.88.[20] This parameter will be important not only in the firms' pricing decision given available information, but also in the firms' attention allocation problem, described below.

[18]See *E-1: Perfect Information*

[19]See *B-2: Optimal Price Setting*

[20]See Mankiw and Reis (2010)

Imperfect Information

Defining the expected value of period profits as

$$\Pi_{jt}(P_{jt}, P_{it}, P_t, Y_t, \varphi_t) \equiv E_{jt}\left[u'(Y_t)\left(P_{jt} - \frac{W_{jt}}{\varphi_{it}}\right)Y_{jt}\right]$$

firm j faces the problem $\max_{P_{jt}} \Pi_{jt}$. Taking a log-quadratic approximation to Π_{jt} around the perfect information non-stochastic equilibrium yields the following formula for firm profits

$$\begin{aligned}\tilde{\Pi}_{jt} =& \hat{\Pi}_1\tilde{p}_{jt} + \hat{\Pi}_{11}\tilde{p}_{jt}^2 + \hat{\Pi}_{12}\tilde{p}_{jt}E_{jt}\tilde{p}_{it} + \hat{\Pi}_{13}\tilde{p}_{jt}E_{jt}\tilde{p}_t + \hat{\Pi}_{14}\tilde{p}_{jt}E_{jt}\tilde{y}_t + \hat{\Pi}_{15}\tilde{p}_{jt}E_{jt}\tilde{\phi}_{it} \\ &+ \text{other terms}\end{aligned}$$

where $\hat{\Pi}_1$ is a constant times the partial derivative of profit with respect to the first argument and the $\hat{\Pi}_{1*}$ coefficients are constants times the second partial derivatives, all evaluated at the point at which all prices are the same; "other terms" collects all terms of the second-order approximation that do not depend on $\tilde{p}_{jt}$ (irrelevent for our purposes since they do not affect the firm's pricing decision).[21] The associated first-order condition yields the following imperfect-information pricing rule

$$\begin{aligned}\tilde{p}_{jt}^* &= -\gamma E_{jt}\tilde{\phi}_{it} + \zeta E_{jt}\tilde{q}_t + (1-\zeta)E_{jt}\tilde{p}_t \\ &= E_{jt}\tilde{p}_{jt}^{\diamond}\end{aligned} \tag{17}$$

To find the imperfect information equilibrium aggregate price we follow a guess and verify approach. Given the form of the perfect-information aggregate price, we guess that under imperfect

[21] See *B-2: Optimal Price Setting* and *C-2: Log-quadratic approximation to an intermediate good firm's profit function* for details.

information it is described by

$$\tilde{p}_t = a\tilde{q}_t - \frac{\gamma}{\zeta}\sum_{l=1}^{I} b_l \mu_l \tilde{\phi}_{lt} \tag{18}$$

This guess will be verified in the next section in conjunction with the solution to the attention allocation problem. In the meantime, substituting this guess in to the imperfect information pricing rule yields

$$\tilde{p}^*_{jt} = \left[(1-\zeta)a + \zeta\right] E_{jt}\tilde{q}_t - (1-\zeta)\frac{\gamma}{\zeta}\sum_{l=1}^{I} b_l \mu_l E_{jt}\tilde{\phi}_{lt} - \gamma E_{jt}\tilde{\phi}_{it} \tag{19}$$

and noticing that since firms only observe Gaussian signals the expected values can be solved using typical signal extraction results:

$$\tilde{p}^*_{jt} = \left[(1-\zeta)a + \zeta\right]\left(\frac{\sigma_q^2}{\sigma_q^2 + \sigma^2_{\psi_j^{(q)}}}\right) s^q_{jt} - (1-\zeta)\frac{\gamma}{\zeta}\sum_{l=1}^{I} b_l \mu_j \left(\frac{\sigma_l^2}{\sigma_l^2 + \sigma^2_{\psi_j^{(l)}}}\right) s^{(l)}_{jt} - \gamma\left(\frac{\sigma_i^2}{\sigma_i^2 + \sigma^2_{\psi_j^{(i)}}}\right) s^{(i)}_{jt}$$

The firm's attention allocation problem is to select optimal signals $s^{(\omega)}_{jt}$. Since the variance of the fundamentals is given, in practice this means that firms will optimally select the variances of the signals' noise.

5.4 Optimal Attention Allocation

Having solved for optimal firm behavior given available information, we now derive the optimal information structure by considering the attention allocation problem.[22] The firm is concerned with the difference between the price it actually sets and the price it would set under full information

[22] See *Appendix D: Information Theory* and *E-2: Rational Inattention under Gaussian White Noise* for all details related to this section.

only to the extent that it results in a loss in profits. This expected loss in profits is

$$E_{jt}\left[\tilde{\Pi}_{jt}\left(\tilde{p}^{\diamond}_{jt},\cdot\right)-\tilde{\Pi}_{jt}\left(\tilde{p}^{*}_{jt},\cdot\right)\right]=\left(\frac{\hat{\Pi}_{11}}{2}\right)E_{jt}\left[\left(\tilde{p}^{\diamond}_{jt}-\tilde{p}^{*}_{jt}\right)^{2}\right]$$

The firm's attention problem then is to optimally select signals to minimize a quadratic loss

$$\min_{\{s^{(\omega)}_{jt}\}_{\omega\in\Omega}} E_{jt}\left[\left(\tilde{p}^{\diamond}_{jt}-\tilde{p}^{*}_{jt}\right)^{2}\right]$$

subject to the constraint that the total information content of the signals does not exceed some value κ. This constraint can be put formally in terms of mutual information

$$\mathcal{I}\left(\left\{\tilde{q}_t,\tilde{\phi}_{1t},\cdots,\tilde{\phi}_{It}\right\};\left\{s^{(q)}_{jt},s^{(1)}_{jt},\cdots,s^{(I)}_{jt}\right\}\right)\leq\kappa$$

Then given the independence assumptions and defining for notational convenience $\kappa^{(\omega)}_j \equiv \mathcal{I}\left(\{\omega_t\};\{s^{(\omega)}_{jt}\}\right)$, it can be reformulated as $\sum_{\omega\in\Omega}\kappa^{(\omega)}_j \leq \kappa$.[23] Since the signals are Gaussian, it can be shown that the mutual information is a function only of the ratio of the variances of the fundamental and the noise

$$\kappa^{(\omega)}_j=\frac{1}{2}\log_2\left(\frac{\sigma^2_\omega}{\sigma^2_{\psi^{(\omega)}_j}}+1\right)\qquad\qquad \omega\in\Omega$$

After some algebra, the firm's attention problem can be finally written

$$\min_{\{\kappa^{(\omega)}_j\}_{\omega\in\Omega}}\sum_{\omega\in\Omega}\left(\bar{\kappa}^{(\omega)}_j\right)^2 2^{-2\kappa^{(\omega)}_j}\qquad;\qquad \bar{\kappa}^{(\omega)}_j=\begin{cases}\left[(1-\zeta)a+\zeta\right]\sigma_q & \omega=q\\ (1-\zeta)\gamma\zeta^{-1}b_l\mu_l\sigma_{\phi_l} & \omega=l\neq i\\ \left[(1-\zeta)\gamma\zeta^{-1}b_i\mu_i+\gamma\right]\sigma_{\phi_i} & \omega=i\end{cases}\tag{20}$$

[23] Recall that ω indexes the set of all stochastic processes Ω - see *Signals* for the definition.

subject to $\sum_{\omega\in\Omega}\kappa_j^{(\omega)}\leq\kappa$ and $\kappa_j^{(\omega)}\geq 0$ for $\omega\in\Omega$. The terms $\bar{\kappa}_j^{(\omega)}$ can be thought of as *importance-weighted volatilities*; their origin is in the firm's optimal imperfect-information pricing rule (19).

This can be solved using standard techniques to yield the following interior solution for the optimal allocation of attention to each fundamental

$$\kappa_j^{(\omega)^*}=\log_2 2^{\bar{\kappa}}+\log_2\bar{\kappa}_j^{(\omega)}-\log_2\bar{\kappa}_j \qquad\qquad \omega\in\Omega \tag{21}$$

where $\bar{\kappa}=\frac{\kappa}{|\Omega|}$ (recall that $|\Omega|$ is the number of stochastic processes) and $\bar{\kappa}_j=\left[\prod_{\omega'\in\Omega}\bar{\kappa}_j^{(\omega')}\right]^{\frac{1}{|\Omega|}}$. This equation is intuitive: the first term gives an equal allocation of attention across all stochastic processes, and the second term adds (subtracts) attention if the importance-weighted volatility of the stochastic process in question is above (below) the (harmonic) mean of importance-weighted volatility across all stochastic processes.

The above is an interior solution; we abuse notation to set

$$\kappa_j^{(\omega)^*}=\begin{cases}\kappa & \kappa_j^{(\omega)^*}>\kappa\\ \kappa_j^{(\omega)^*} & \kappa_j^{(\omega)^*}\in[0,\kappa]\\ 0 & \kappa_j^{(\omega)^*}<0\end{cases}$$

To fully incorporate corner solutions with more than two options we need to specify a solution process in the case that $\kappa_j^{(\omega)}<0$, which corresponds to a "negative" attention allocation to some stochastic process. This is not allowed. When this occurs, the problem can be solved by replacing the negative value with zero and then rescaling the remaining allocations such that they sum to κ.

Using these optimal attention allocations in the firms' imperfect-information pricing rules and integrating across all firms yields the imperfect-information equilibrium aggregate price

$$\tilde{p}_t^* = \underbrace{\left[\sum_{i=1}^{I} \mu_l[(1-\zeta)a+\zeta]\left(1-2^{-2\kappa_i^{(q)*}}\right)\right]}_{a} \tilde{q}_t - \frac{\gamma}{\zeta}\sum_{l=1}^{I}\underbrace{\left[\sum_{i=1}^{I} w_{li}\left(1-2^{-2\kappa_i^{(l)*}}\right)\right]}_{b_l} \mu_l\tilde{\phi}_{lt} \tag{22}$$

where $w_{li} = \big[(1-\zeta)b_l\mu_i + \zeta\mathbf{1}(l=i)\big]$ and $\mathbf{1}(l=i)$ is the indicator function that takes the value 1 if $l = i$ and is 0 otherwise. We have applied symmetry across within-industry firms to write $\kappa_j^{(\omega)*} = \kappa_i^{(\omega)*}\ j \in J_i,\ i = 1, \ldots, I$. This verifies our guess.

Notice that through the term $\bar{\kappa}_j$, the optimal attention allocation for each fundamental depends on all of the coefficients a, $\{b_l\}_{l=1}^{I}$. For this reason, in general there are not analytic solutions to these coefficients, although computational techniques can be used to solve the fixed point problem.

6 Results

6.1 Interpretation

First we will consider the terms $\bar{\kappa}_j^{(\omega)}$ from the attention allocation objective function. They can be written as

$$\bar{\kappa}_j^{(\omega)} = \begin{cases} (1-\zeta)a\sigma_q + \zeta\sigma_q & \omega = q \\ (1-\zeta)\frac{\gamma}{\zeta}b_l\mu_l\sigma_{\phi_l} & \omega = l \neq i \\ (1-\zeta)\frac{\gamma}{\zeta}b_i\mu_i\sigma_{\phi_i} + \gamma\sigma_{\phi_i} & \omega = i \end{cases} \tag{23}$$

They are the products of structural parameters related to the importance of the process to the firm's pricing decision along with the parameter governing the volatility of the shock, and so may be considered as terms describing "importance-weighted volatility". In that light, the firm's objective is to minimize the overall importance-weighted volatility that they face.

It can be seen, moreover, that the importance of the process is derived from two distinct mechanisms related to the firm's pricing problem. First, there are direct effects: nominal aggregate demand has a direct effect on the demand curve for the firm's differentiated product and the firm's sector-specific shock has a direct effect on marginal costs. These direct effects appear as the terms that do not include the strategic complementarity $(1-\zeta)$ and do not depend on the coefficients that determine the aggregate price level $a, \{b_l\}_{l=1}^{I}$.

The second type of effect is due to strategic complementarities that arise insofar as firms are concerned with their *relative* price. If $(1-\zeta) > 0$, so that there are strategic complementarities, firms will want to raise their price in response to a general increase in prices. Larger coefficients $a, \{b_l\}_{l=1}^{I}$ amplify the effect of strategic complementarities through a larger response of the aggregate price level to firms' responses to the direct effects described above.

Finally, notice that the problem is complicated by the circularity of definitions: the firms' objective depends on the coefficients determining the aggregate price, which in turn depend on the solution to the firms' objective problem. Thus the coefficients and optimal attention levels are determined by the solution to a fixed point problem.

The optimal allocation of attention to a stochastic process can be written

$$\kappa_j^{(\omega)} = \frac{\kappa}{|\Omega|} + \log_2\left(\frac{\bar{\kappa}_j^{(\omega)}}{\left[\prod_{\omega' \in \Omega} \bar{\kappa}_j^{(\omega')}\right]^{\frac{1}{|\Omega|}}}\right) \tag{24}$$

where κ denotes the total information processing capacity available to the agent, $|\Omega|$ is the total number of stochastic processes and the $\bar{\kappa}_j^{(\omega)}$ terms are as just described. The first observation is that if all importance-weighted volatility terms were equal, the second term would be zero and the optimal attention allocation would be to evenly divide capacity across all shocks. The second observation is that the denominator of the second term can be interpreted as an "average" importance-weighted volatility across all shocks. Then the optimal allocation gives more attention to those processes whose importance-weighted volatility exceeds the "average" and less attention to those that fall below the “average”. Finally, notice that through the first term of the optimal allocation, as total capacity becomes infinitely large, the level of attention devoted to each shock also becomes infinitely large.

Returning to the coefficients induced by the firms’ optimal allocations, notice that they can be rewritten to emphasize the strategic complementarities parameter

$$
\begin{aligned}
a &= (1-\zeta)a \sum_{i=1}^{I} \mu_i \left(1 - 2^{-2\kappa_i^{(q)*}}\right) + \zeta \sum_{i=1}^{I} \mu_i \left(1 - 2^{-2\kappa_i^{(q)*}}\right) \\
b_l &= (1-\zeta)b_l \sum_{i=1}^{I} \mu_i \left(1 - 2^{-2\kappa_i^{(l)*}}\right) + \zeta \left(1 - 2^{-2\kappa_l^{(l)*}}\right)
\end{aligned}
$$

As noted above, this exposes the constitution of these coefficients as combinations - weighted by the strategic complementarities parameter - of the average direct effect of shocks to firms, weighted by sector size, and the effect arising through the influence of the aggregate price level on firms’ *relative* prices. One interesting implication is that as strategic complementarities become strong, so that $(1-\zeta) \to 1$, the coefficient a governing the influence of monetary policy on the aggregate price level is decoupled from real aggregate demand, since that term disappears from firms’ pricing rules.

Note that as total capacity tends to infinity, so that the model tends to perfectly informed agents, we have

$$
\begin{aligned}
a &\to (1-\zeta)a + \zeta \qquad &\implies \qquad & a \to 1 \\
b_l &\to (1-\zeta)b_l + \zeta \qquad &\implies \qquad & b_l \to 1
\end{aligned}
$$

Thus the aggregate price under imperfect information tends to the aggregate price under perfect information as total information processing capacity becomes arbitrarily large.

6.2 A one-sector model

In the special case of a one-sector model, we have $I = 1$, $b_l \equiv 0$, and $\mu_1 = 1$. The imperfect information equilibrium aggregate price level is then

$$\tilde{p}_t^* = \underbrace{[(1-\zeta)a + \zeta]\left(1 - 2^{-2\kappa_1^{(q)^*}}\right)}_{a} \tilde{q}_t$$

and in this case, the coefficient a can be solved for explicitly

$$a = \begin{cases} \frac{(2^{2\kappa}-1)\zeta}{1+(2^{2\kappa}-1)\zeta} & \kappa_1^{(q)^*} > \kappa \\ 1 - 2^{-\kappa}\left(\frac{\gamma}{\zeta}\right)\left(\frac{\sigma_{\phi_1}}{\sigma_q}\right) & \kappa_1^{(q)^*} \in [0, \kappa] \\ 0 & \kappa_1^{(q)^*} < 0 \end{cases}$$

This is identical to the result in MW section IV, and although it appears in a slightly different form, the interpretation has the same implications as the more general discussion above.

Notice that here there is no transmission mechanism for idiosyncratic shocks to affect the aggregate price level. Although this is a one-sector model, so that the firm literally has only two signals to observe (their own productivity shock and monetary policy), the multisector extension in MWW does not depart too far from this approach in that firms still receive two signals, one regarding aggregate conditions and one regarding idiosyncratic conditions. The above discussion of the current paper's results demonstrates that there are important subtleties that arise from idiosyncratic and aggregate components of each shock, re-emphasizing the previous arguments in favor of the current paper's approach, which models each firm's attention allocation problem between each stochastic shock separately.

6.3 Calibrating volatilities

The basic point of the rational inattention approach is that firms may not immediately react to monetary policy (or other aggregate shocks) if they are not paying attention to them. Maćkowiak and Wiederholt (2009) formalized this concept and show that when the volatility of aggregate shocks is low relative to the volatility of a sector-specific shock, firms will optimally devote more of their attention to sector-specific conditions. If the volatility differential is great enough, the lack of attention paid to aggregate shocks will imply slow adjustment, meaning that prices will appear to be sticky in response to aggregate shocks. In their model, to achieve the appropriate degree of stickiness, they suggest that the differential in standard deviations needs to be an order of magnitude.

In particular, they calibrate idiosyncratic volatilities to match the size of average absolute price changes; this is performed under perfect information. Thus the term calibrated is the absolute expected value of $\tilde{p}^{\diamond}_{jt}$. When all disturbances are Gaussian, the absolute value is distributed half-normal with expected value $E[|\tilde{p}^{\diamond}_{jt}|] = \sigma_{p_j}\sqrt{\frac{2}{\pi}}$. Equilibrium in the one-sector and multi-sector models, along with independence assumptions, implies

$$\sigma^2_{p_j} = \sigma^2_q + \gamma^2\sigma^2_{\phi_i} \qquad \text{One-sector}$$

$$\sigma^2_{p_j} = \sigma^2_q + \gamma^2\sigma^2_{\phi_i} + \left(\frac{1-\zeta}{\zeta}\right)^2 \gamma^2 \sum_{l=1}^{I} \mu_l^2 \sigma^2_{\phi_l} \qquad \text{Multi-sector}$$

In the baseline model calibrated in MW, γ is normalized to one since changes to it have the same practical effect on the model as changes to $\sigma^2_{\phi_i}$. The calibration exercise fixes the value of σ^2_q according to the observed volatility of detrended nominal GNP, and fixes the value of $\sigma^2_{p_j}$ according to the size of average absolute price changes, as described above. Thus the volatility of the idiosyncratic shocks is fixed by their difference. The resultant calibration then has $\sigma_z = 11.8\sigma_q$.

However, empirical work suggests that while the aggregate component is less volatile than the

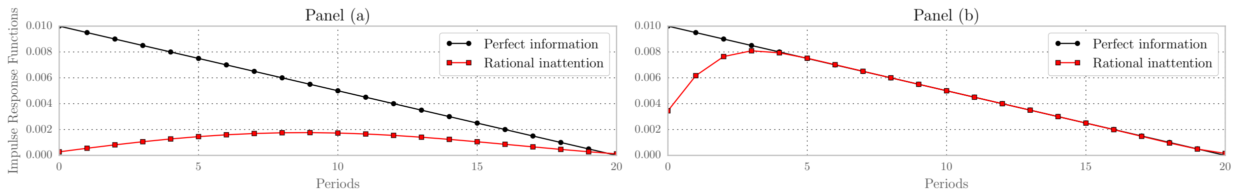

Fig. 1: Impulse response functions in the one-sector model. Panel (a) shows the impulse responses when $\sigma_z = 11.8\sigma_q$ (as in Maćkowiak and Wiederholt, 2009) and Panel (b) shows the impulse responses when $\sigma_z = 3.03\sigma_q$ (as suggested by results from Boivin et al., 2009)

idiosyncratic component, it is not an order of magnitude less volatile; for example, the decomposition in Boivin et al. (2009) find that for the average firm, the standard deviation of the common component (σ_q, above) is 0.33 while the standard deviation of the aggregate component (σ_z, above) is 1.09. For the median it is even closer, at 0.27 and 0.71 respectively. Even with the average values, this suggests that $\sigma_z = 3.03\sigma_q$. The effect of this change on the model is illustrated in Fig. 1.[24]

Furthermore, as they point out, their calibration is conservative in a number of ways. First, they exclude sales, which lowers the reported average absolute price change from 11.5% to 9.5%, and second they calibrate against the perfect information equilibrium rather than the rational inattention equilibrium. Changing either of these decisions would force their calibration to yield even more idiosyncratic volatility.

Finally, as developed in this paper, it is not merely the volatility of shocks that matters, it is also the weight in the pricing solution. This insight is important for calibrations. Here, MW normalize $\gamma = 1$, so what they calibrate as σ_z is actually $\gamma\sigma_z$. Now, if $\gamma\sigma_z = 11.8\sigma_q$, then $\sigma_z = \frac{11.8}{\gamma}\sigma_q$. Recalling that $\gamma = (1+\varepsilon)/(1+\rho\varepsilon)$ and using calibrations as in Mankiw and Reis, 2010, we calculate $\gamma \approx 1/5$, so that $\frac{1}{\gamma} \approx 5$ (alternative reasonable calibrations make γ even smaller). Thus their conservative calibration actually requires $\sigma_z = 59\sigma_q$. This is far removed from the results of Boivin et al. (2009).

[24]These figures were created using the MATLAB programs accompanying Maćkowiak and Wiederholt (2009) by varying the volatility calibration.

What this analysis points out is not that the approach of Maćkowiak and Wiederholt (2009) is fundamentally flawed, but rather that a more complex model is required in order to achieve realistic calibrations. Their insight that aggregate demand (i.e. monetary) shocks play a relatively small role in firms' pricing decisions merely requires that, in the language introduced above, the importance weighted volatility of aggregate demand shocks be small relative to that of other shocks.

Now the intuition for why a multisector approach is appealing is easy to see. To achieve aggregate price stickiness, all that is required is that firms pay little attention to aggregate conditions. Whereas in the one-sector case there were only two types of conditions to pay attention to (so that a decrease in attention to one meant an increase in attention to the other), here firms also pay attention to each other. This creates two channels by which an increase in idiosyncratic volatility reduces attention paid to aggregate conditions. First, firms are still concerned about their own productivity shock, which influences their marginal costs directly, and second they are concerned about the productivity shocks to all other firms because of the general equilibrium effect on aggregate demand and the aggregate price level.

In essence, the multisector model gives firms more reasons to pay attention to idiosyncratic components, which leaves less attention available for the aggregate component. The the multisector model increases the *importance* of idiosyncratic shocks, even when *volatilities* are constant. Since the firm's attention allocation decision depends on *importance-weighted volatilities*, this model reduces the required volatility differential by increasing the importance differential.[25]

7 Extension: relative demand shocks and intermediate inputs

In this section, we consider augmenting household and firm behavior to include relative demand shocks, composite productivity shocks, and intermediate inputs. These variations not only intro-

[25]For details of the importance-weighted volatilities, see section *Interpretation*.

duce desirable model characteristics but also provide additional motivation the implicit claim in the baseline model that firms pay attention to each other. Here, firms must pay attention to each other because their production process requires intermediate inputs.[26] To introduce demand shocks, we replace the demand weight μ_i with $D_{it}\mu_i$ so that the nested CES Dixit-Stiglitz aggregators can be written

$$C_{hit} = \left[\int_{J_i} (D_{it}\mu_i)^{r-1} C_{hjt}^r dj\right]^{\frac{1}{r}} \qquad C_{ht} = \left[\sum_{i=1}^{I} (D_{it}\mu_i)^{1-p} C_{hit}^p\right]^{\frac{1}{p}}$$

and where every period we require $\sum_i D_{it}\mu_i = 1$.[27] All other changes result from modifications to firms' production functions. In particular, we write:

$$Y_{jt} = \Phi_{jt} n_{jt}^{\alpha} X_{jt}^{1-\alpha}$$

where $\Phi_{ijt} = \varphi_t \varphi_{it} \varphi_{jt}$ is the composite productivity shock, and X_{jt} is a composite intermediate input constructed from the output of other firms. In particular, we suppose that, like the demand composite, it exhibits constant elasticity of substitution, so that

$$X_{jt} = \left[\sum_{k=1}^{I} (D_{kt}\mu_k)^{1-p} X_{ijkt}^p\right]^{1/p} \qquad X_{ijkt} = \left[\int_{J_k} (D_{kt}\mu_k)^{r-1} X_{ijklt}^r dl\right]^{1/r}$$

where X_{ijklt} is the quantity of the good produced by firm l (in sector k) used by firm j (in sector i). Similarly, X_{ijkt} is a composite of all the goods produced by firms in sector k used by firm j (in sector i). Finally, X_{jt} is a composite of the goods produced by all firms used by firm j.

Along the same lines as the baseline model, it is not too hard to show that this results in the

[26] See Basu (1995), Bouakez et al. (2009), and Carvalho and Lee (2011) for examples of similar models with Calvo-type pricing.

[27] Notice that this specification nests the baseline model when $D_{it} \equiv 1, i = 1, \ldots, I$.

following log-linear pricing equation[28]

$$\tilde{p}^{\diamond}_{jt} = \underline{\psi}\tilde{d}_{it} - \underline{\upsilon}\tilde{\phi}_{jt} - \underline{\gamma}(\tilde{\phi}_t + \tilde{\phi}_{it}) + \underline{\zeta}\tilde{q}_t + (1 - \underline{\zeta})\tilde{p}_t$$

where aggregate prices evolve according to

$$\tilde{p}_t = \tilde{q}_t - \frac{\underline{\gamma}}{\underline{\zeta}}\sum_{i=1}^{I}\mu_i\tilde{\phi}_{it} - \frac{\underline{\gamma}}{\underline{\zeta}}\tilde{\phi}_t$$

These equations differ from those in the baseline model through more complex parameters and the introduction of new shocks ($d_{it} \equiv \log D_{it}$ is the demand shock, and ϕ_{jt}, ϕ_t are the new productivity shock components). Qualitatively, however, it tells much the same story. Under reasonable calibrations, all of the parameters above are positive, indicating that the price set by firm j rises with relative and aggregate demand, falls with increased productivity (either firm-level, sectoral, or aggregate), and, as long as there are strategic complementarities, increases with the aggregate price level.

Similarly, the rational inattention solution is qualitatively the same. More attention will be paid to those shocks which are relatively more important (the coefficient in the above pricing equation is higher) or more volatile. The log-linear rational inattention price-setting problem reduces to a sum of weighted shocks, which can be solved in the white noise case as described above, and again the unknown coefficients can be found as the solution to a fixed-point problem. Finally, it preserves the calibration result introduced above, that through the effect of multiple targets for a firm's attention prices can adjust slowly with respect to monetary policy shocks while responding quickly to idiosyncratic shocks, without requiring unrealistic volatility differentials.

[28]For example see the equation for marginal costs in Carvalho and Lee (2011).

8 Conclusion

In this paper we extend the rational inattention model of price-setting to account for multiple sectors in which firms care about their own idiosyncratic shocks, idiosyncratic shocks to other firms, and aggregate shocks. In addition to the baseline model, we consider an extension including relative demand shocks and intermediate inputs. We derive optimal attention allocations and the implied optimal price-setting behavior, allowing us to consider the effect of various parameterizations on the responsiveness of prices. The functional forms derived herein inform new directions for continued empirical research into price-setting behavior.

Our results provide several novel contributions. First, we allow firms to exhibit heterogenous behavior that depends not only on their own idiosyncratic shocks but also on firm characteristics and their relationship to other firms. Second, we show that not only can the model generate slow responses to aggregate shocks along with quick responses to idiosyncratic shocks, it can do so with less extreme parameter calibrations than in related work. Finally, we emphasize the role of importance-weighted volatility in generating optimal attention allocations, rather than volatility only.

Finally, the basic model considered here provides a baseline for future research. It would be interesting, for example, to further extend the multi-sector model to account for additional firm characteristics in order to derive testable cross-sectional implications for price-stickiness; another interesting direction is to introduce network effects as in Acemoglu et al. (2012).

Appendix A: Model

8.1 A-1: Constant Elasticity of Substitution Preferences

A-1.1: Definition: Consumption good

The composite consumption good is defined as a monotonic transformation of the generalized mean $\tilde{C}_{ht}$ as follows:

$$\tilde{C}_{ht} = \left[\frac{\sum_{i=1}^{I} \mu_i^{1-p} C_{hit}^p}{\sum_{i=1}^{I} \mu_i^{1-p}} \right]^{\frac{1}{p}}$$

$$C_{ht} = \left[\sum_{i=1}^{I} \mu_i^{1-p} C_{hit}^p \right]^{\frac{1}{p}} = \left[\sum_{i=1}^{I} \mu_i^{1-p} \right]^{\frac{1}{p}} \tilde{C}_{ht}$$

The exponent on the weight term is a normalization so that the resulting price index has the property that if every industry-level good has the same price, that price also is the index price. Furthermore, if all prices are the same then the derived demand for each industry-level good is just the fraction of the demand for the composite good weighted by the industry's size. Mathematically $C_{hit}^d = \mu_i C_{ht}$, and since we normalized the total measure of goods to one, $\mu_i \in [0, 1]$ for each industry i. This approach is the same as in Woodford (2003).

A-1.2: Definition: Industry-level good

The composite industry-level good is defined as a monotonic transformation of the generalized mean $\tilde{C}_{hit}$ as follows:

$$\tilde{C}_{hit} = \left[\frac{\int_{J_i} C_{hjt}^r dj}{\int_{J_i} 1 dj} \right]^{\frac{1}{r}} = \left[\int_{J_i} \mu_i^{-1} C_{hjt}^r dj \right]^{\frac{1}{r}}$$

$$C_{hit} = \left[\int_{J_i} \mu_i^{r-1} C_{hjt}^r \right]^{\frac{1}{r}} = \left[\mu_i^r \int_{J_i} \mu_i^{-1} C_{hjt}^r \right]^{\frac{1}{r}} = \mu_i \tilde{C}_{hit}$$

The exponent on the weight term is for the same normalizing purpose as above.

A-1.3: Demand: Constant Elasticity of Substitution Preferences

As in Dixit and Stiglitz (1977), we can use a multi-stage budgeting procedure to first solve for the demand for industry-level and intermediate goods' demand in terms of the consumers' total demand for the consumption good, and then solve their inter-temporal problem in terms only of the consumption good.

The first stage is itself split into two steps: (1) solve for industry-level demand in terms of total demand, and (2) solve for intermediate good demand in terms of industry-level demand.

Step 1: Industry-level demand

The interpretation of the definition of the consumption good is as a *utility specification*. Thus solving for demand is the standard microeconomic constrained optimization problem

$$\max_{\{C_{hit}\}_{i=1}^{I}} u\left(\{C_i\}_{i=1}^{I}\right)$$

subject to $\sum_{i=1}^{I} C_{hit} P_i = W$, where P_{it} is the price of industry-level good i at time t and W is total wealth, and where the utility specification is the generalized mean, above:

$$u\left(\{C_i\}_{i=1}^{I}\right) = \tilde{C}_{ht} = \left[\frac{\sum_{i=1}^{I} \mu_i^{1-p} C_{hit}^{p}}{\sum_{i=1}^{I} \mu_i^{1-p}}\right]^{\frac{1}{p}}$$

The model used in the paper is a monotonic transformation of this specification, and it will yield equivalent demand specifications due the ordinal nature of utility functions.

This constrained optimization problem can be solved by forming a Lagrangian and taking first-order conditions. To ease notation, define $w_i \equiv \frac{\mu_i^{1-p}}{\sum_{i=1}^{I} \mu_i^{1-p}}$.

$$\mathcal{L} = \left[\sum_{i=1}^{I} w_i C_{hit}^{p}\right]^{\frac{1}{p}} - \lambda \left[\sum_{i=1}^{I} C_{hit} P_{it} - W\right]$$

Assuming an interior solution, the I first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial C_{hit}} = 0 = \frac{1}{p}\left[\sum_{i=1}^{I} w_i C_{hit}^{p}\right]^{\frac{1}{p}-1} w_i p C_{hit}^{p-1} - \lambda P_{it} = w_i u\left(\{C_i\}_{i=1}^{I}\right)^{1-p} C_{hit}^{p-1} - \lambda P_{it}$$

$$C_{hit} = \left(\frac{\lambda P_{it}}{w_i}\right)^{\frac{1}{p-1}} \tilde{C}_{ht}$$

This yields the demand for the industry-level good. The Lagrangian multiplier λ is the marginal value of relaxing the constraint, or the marginal value of wealth.

$$\tilde{C}_{ht} = \left[\sum_{i=1}^{I} w_i C_{hit}^{p}\right]^{\frac{1}{p}} = \left[\sum_{i=1}^{I} w_i \left(\left(\frac{\lambda P_{it}}{w_i}\right)^{\frac{1}{p-1}} \tilde{C}_{ht}\right)^{p}\right]^{\frac{1}{p}}$$

$$= \tilde{C}_{ht} \lambda^{\frac{1}{p-1}} \left[\sum_{i=1}^{I} w_i^{\frac{1}{1-p}} P_{it}^{\frac{p}{p-1}}\right]^{\frac{1}{p}}$$

$$\frac{1}{\lambda} = \left[\sum_{i=1}^{I} w_i^{\frac{1}{1-p}} P_{it}^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$$

The price index is the price of the composite good $\tilde{C}_{ht}$, which is equivalently the price of increasing utility. This quantity is the inverse of the marginal value of wealth, so that

$$P_t \equiv \frac{1}{\lambda} = \left[\sum_{i=1}^{I} w_i^{\frac{1}{1-p}} P_{it}^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$$

Notice that if all industry-level prices are the same, so that $P_{it} = P_{i't} = \bar{P}_t$, then:

$$P_t = \left[\sum_{i=1}^{I} w_i^{\frac{1}{1-p}} \bar{P}_t^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$$
$$= \bar{P}_t \left[\sum_{i=1}^{I} w_i^{\frac{1}{1-p}}\right]^{\frac{p-1}{p}}$$

Thus if we want to have the property that in this case $P_t = \bar{P}_t$, then we must have $\left[\sum_{i=1}^{I} w_i^{\frac{1}{1-p}}\right]^{\frac{p-1}{p}} = 1$. This does not hold for $\tilde{C}_{ht}$, but it does hold for the transformation C_{ht} since in that case $w_i \equiv \mu_i^{1-p}$ and then

$$\left[\sum_{i=1}^{I} w_i^{\frac{1}{1-p}}\right]^{\frac{p-1}{p}} = \left[\sum_{i=1}^{I} \mu_i^{\frac{1-p}{1-p}}\right]^{\frac{p-1}{p}} = 1^{\frac{p-1}{p}} = 1$$

Finally we can rewrite industry-level demand

$$C_{hit} = w_i^{\frac{1}{1-p}} \left(\frac{P_{it}}{P_t}\right)^{\frac{1}{p-1}} \tilde{C}_{ht}$$

If we use the transformed C_{ht}, then this reduces to:

$$C_{hit} = \mu_i^{\frac{1-p}{1-p}} \left(\frac{P_{it}}{P_t}\right)^{\frac{1}{p-1}} C_{ht}$$

Collecting the final demand function and price index for the transformed C_{ht}, we have

$$C_{hit} = \mu_i \left(\frac{P_{it}}{P_t}\right)^{\frac{1}{p-1}} C_{ht}$$

$$P_t = \left[\sum_{i=1}^{I} \mu_i P_{it}^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$$

Step 2: Intermediate good demand

Following similar steps as above, the final demand function and price index are given by

$$C_{hjt} = \frac{1}{\mu_i}\left(\frac{P_{jt}}{P_{it}}\right)^{\frac{1}{r-1}} C_{hit}$$

$$P_{it} = \left[\int_{J_i} \frac{1}{\mu_i} P_{jt}^{\frac{r}{r-1}} dj\right]^{\frac{r-1}{r}}$$

And the CES demand function for intermediate goods in terms of the consumption good is

$$\begin{aligned} C_{hjt} &= \frac{1}{\mu_i}\left(\frac{P_{jt}}{P_{it}}\right)^{\frac{1}{r-1}} \mu_i \left(\frac{P_{it}}{P_t}\right)^{\frac{1}{p-1}} C_{ht} \\ &= P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{1}{1-r}+\frac{1}{p-1}} P_t^{\frac{1}{1-p}} C_{ht} \end{aligned}$$

8.2 A-2: Budget Contraints

In each period, households purchase (1) consumption goods, and (2) invest in risk-free bonds. They receive income from (1) wages, (2) a share of intermediate goods firm profits, and (3) investment income from bonds purchased in the previous period.

Assume that all households are endowed with equal ownership shares in each of the intermediate goods firms. Then each household's share of the profits can be denoted π_{jt}.

Bonds are indexed by time period in which they mature so that B_t refers to bonds purchased in time $t-1$ that yield income in period t. The gross nominal rate of return on a bond purchased in period $t-1$ is denoted R_t. The bonds are riskless, so that R_t is known in period $t-1$.

The nominal flow budget constraint is

$$\int_J P_{jt}C_{hjt}dj + B_{ht+1} \leq \int_J \theta_{hj}\pi_{jt}dj + \int_J W_{jt}n_{hjt}dj + R_tB_{ht}$$

Plugging the CES demand functions derived above into the consumption spending portion of the budgent constraint yiels

$$\begin{aligned}
\int_J P_{jt}P_{jt}^{\frac{1}{r-1}}P_{it}^{\frac{1}{1-r}+\frac{1}{p-1}}P_t^{\frac{1}{1-p}}C_{ht}dj &= P_t^{\frac{1}{1-p}}C_{ht}\int_J P_{jt}^{\frac{r}{r-1}}P_{it}^{\frac{1}{1-r}+\frac{1}{p-1}}dj \\
&= P_t^{\frac{1}{1-p}}C_{ht}\sum_{i=1}^{I}\mu_i P_{it}^{\frac{r}{r-1}}P_{it}^{\frac{1}{1-r}} \\
&= P_t^{\frac{1}{1-p}}C_{ht}\sum_{i=1}^{I}\mu_i P_{it}^{\frac{p}{p-1}} \\
&= P_t^{\frac{1}{1-p}}C_{ht}P_t^{\frac{p}{p-1}} \\
&= P_tC_{ht}
\end{aligned}$$

Using this, the nominal flow budget constraint can be rewritten

$$P_tC_{ht} + B_{ht+1} \leq \int_J \theta_{hj}\pi_{jt}dj + \int_J W_{jt}n_{hjt}dj + R_tB_{ht}$$

Appendix B: Optimal Behavior

8.3 B-1: Optimal Household Behavior

B-1.1: Sequential Problem

The representative household's problem is

$$\max_{\{C_t\}_{t\geq 0}\{n_{jt}\}_{j\in J,t\geq 0}} = E_0 \sum_{t=0}^{\infty} \beta^t \left[u(C_t) - \int_J v(n_{jt}) dj \right]$$

subject to the nominal budget constraint

$$P_t C_t + B_{t+1} \leq \int_J \pi_{jt} dj + \int_J W_{jt} n_{jt} dj + R_t B_t$$

Define wealth at time t as

$$A_t = \int_J \pi_{jt} dj + \int_J W_{jt} n_{jt} dj + R_t B_t$$

Notice that given wealth and the household's consumption choice, bond holdings are determined by $B_{t+1} = A_t - P_t C_t$.

B-1.2: Bellman system

The solution to the sequential problem is equivalent to the solution to the following functional equation

$$V(A) = \max_{C,\{n_j\}_{j\in J}} \left\{ u(C) - \int_J v(n_j) dj + \beta E \left[V(A') \right] \right\}$$

subject to

$$
\begin{aligned}
A' &= \int_J \pi_j' dj + \int_J W_j' n_j' dj + R'B' \\
&= \int_J \pi_j' dj + \int_J W_j' n_j' dj + R'(A - PC)
\end{aligned}
$$

First-order Conditions

$$
\begin{aligned}
0 &= \frac{\partial V(A)}{\partial C} = u'(C) + \beta E\left[V'(A')\right](-P)R' \\
0 &= \frac{\partial V(A)}{\partial n_j} = -v'(n_j) + \beta V'(A') W_j R'
\end{aligned}
$$

Envelope Condition

$$
V'(A) = \beta V'(A')R'
$$

Euler Equation

Combining the first-order condition for consumption and the envelope condition yields

$$
V'(A) = \frac{u'(C)}{P}
$$

which can then be forwarded and plugged back into the first-order condition for consumption to give the household's Euler equation governing intertemporal consumption tradeoffs

$$u'(C_t) = \beta E_t \left[R_{t+1} \frac{P_t}{P_{t+1}} u'(C_{t+1}) \right]$$

Static First-order Condition

Then from the first-order condition for labor we get

$$v'(n_{jt}) = \frac{W_{jt}}{P_t} u'(C_t)$$

8.4 B-2: Optimal Price Setting

B-2.1: Perfect Information

Firms face the problem

$$\max_{P_{jt}} u'(C_t) \left(P_{jt} - \frac{W_{jt}}{\varphi_{it}} \right) Y_{jt}$$

which can be rewritten using the CES demand function as

$$\max_{P_{jt}} u'(C_t) \left(P_{jt} - \frac{W_{jt}}{\varphi_{it}} \right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{1}{1-r} + \frac{1}{p-1}} P_t^{\frac{1}{1-p}} C_t$$

Their first-order condition is:

$$
\begin{aligned}
0 &= u'(C_t)\left[\left(1+\frac{1}{r-1}\right)P_{jt}^{\frac{1}{r-1}} - \left(\frac{1}{r-1}\frac{W_{jt}}{\varphi_{it}}\right)P_{jt}^{\frac{1}{r-1}-1}\right]P_{it}^{\frac{1}{1-r}+\frac{1}{p-1}}P_t^{\frac{1}{1-p}}C_t \\
rP_{jt}^{\frac{1}{r-1}} &= \frac{W_{jt}}{\varphi_{it}}P_{jt}^{\frac{1}{r-1}-1} \\
P_{jt} &= \frac{1}{r}\frac{W_{jt}}{\varphi_{it}}
\end{aligned}
$$

This is the standard result that monopolists set price as a markup over marginal costs.

Proceed by substituting out wages using the household's static first-order condition and using (1) the goods market clearing condition, (2) the production function, and (3) the demand function for the intermediate good

$$
\begin{aligned}
P_{jt} &= \frac{1}{r}\frac{1}{\varphi_{it}}\left[P_t\frac{n^{\varepsilon}}{C_t^{-\sigma}}\right] \\
&= \frac{1}{r}\frac{1}{\varphi_{it}}\left[P_tY_t^{\sigma}\left(\frac{Y_{jt}}{\varphi_{it}}\right)^{\varepsilon}\right] \\
&= \frac{1}{r}\left(\frac{1}{\varphi_{it}}\right)^{1+\varepsilon}\left[P_tY_t^{\sigma}\left(P_{jt}^{\frac{1}{r-1}}P_{it}^{\frac{1}{1-r}+\frac{1}{p-1}}P_t^{\frac{1}{1-p}}Y_t\right)^{\varepsilon}\right]
\end{aligned}
$$

Since productivity shocks are industry-level and they represent the only difference between firms, we can now apply symmetry between all firms within a given industry to note that $P_{jt} = P_{it}$.

$$
\begin{aligned}
P_{jt} &= \frac{1}{r}\left(\frac{1}{\varphi_{it}}\right)^{1+\varepsilon}\left[P_{jt}^{\frac{\varepsilon}{p-1}}P_t^{\frac{1-p+\varepsilon}{1-p}}Y_t^{\sigma+\varepsilon}\right] \\
P_{jt}^{\frac{1-p+\varepsilon}{1-p}} &= \left[\frac{1}{r}\left(\frac{1}{\varphi_{it}}\right)^{1+\varepsilon}Y_t^{\sigma+\varepsilon}\right]P_t^{\frac{1-p+\varepsilon}{1-p}} \\
P_{jt} &= \left[\frac{1}{r}\left(\frac{1}{\varphi_{it}}\right)^{1+\varepsilon}Y_t^{\sigma+\varepsilon}\right]^{\frac{1-p}{1-p+\varepsilon}}P_t^{\frac{1-p+\varepsilon}{1-p+\varepsilon}}
\end{aligned}
$$

Defining $\alpha \equiv \frac{1-p}{1-p+\varepsilon} = \frac{1}{1+\rho\varepsilon}$ we arrive at the final perfect information pricing equation

$$P^{\diamond}_{jt} = \left[\frac{1}{r}\left(\frac{1}{\varphi_{it}}\right)^{1+\varepsilon} Y_t^{\sigma+\varepsilon}\right]^{\alpha} P_t$$

It will also be convenient to have this expression with variables in log-form, where lowercase variables denote logs of uppercase variables

$$p^{\diamond}_{jt} = \alpha \log\frac{1}{r} - \alpha(1+\varepsilon)\phi_{it} + \alpha(\sigma+\varepsilon)y_t + p_t$$

To expose strategic complementarities define $\zeta = \alpha(\sigma+\varepsilon)$, and for notational convenience define $\gamma = \alpha(1+\varepsilon)$. Recall also that $q_t = p_t + y_t$. Then the firms' perfect information pricing rule is

$$p^{\diamond}_{jt} = \alpha \log\frac{1}{r} - \gamma\phi_{it} + \zeta q_t + (1-\zeta)p_t$$

To aid interpretation of the imperfect information pricing rule we express the pricing-rule in proportional deviation from common price form below by defining $\tilde{x}_t \equiv \frac{(X_t - \bar{X})}{\bar{X}} \approx \log\left(\frac{X_t}{\bar{X}}\right) = x_t - \bar{x}$

$$p^{\diamond}_{jt} - \bar{p}_{jt} = \left(\alpha\log\frac{1}{r} - \gamma\phi_{it} + \zeta q_t + (1-\zeta)p_t\right) - \left(\alpha\log\frac{1}{r} - \gamma\bar{\phi}_{it} + \zeta\bar{q}_t + (1-\zeta)\bar{p}\right)$$

which reduces to

$$\tilde{p}^{\diamond}_{jt} = -\gamma\tilde{\phi}_{it} + \zeta\tilde{q}_t + (1-\zeta)\tilde{p}_t$$

B-2.2: Imperfect Information

This section describes optimal firm behavior under imperfect information using results based on log approximations; it is just a summary of the approximations derived in detail in *Appendix C:*

Log Approximations.

The generic problem facing an intermediate goods firm is given above in (8.4). In equilibrium, the objective can be written as a function only of prices, shocks, and aggregate output (this is because wages can be substituted out as a function of these variables, using the households' static first-order condition)

$$\Pi_{jt}(P_{jt}, P_{it}, P_t, Y_t, \varphi_t)$$

Given this, the firm's problem can be expressed as

$$\max_{P_{jt}} \Pi_{jt}$$

We proceed by taking a log-quadratic approximation to Π_{jt} around the perfect-information equilibrium,

$$\begin{aligned}\tilde{\Pi}_{jt} =& \Pi_1 \bar{P}\tilde{p}_{jt} + \frac{\Pi_{11}}{2!}\bar{P}^2\tilde{p}_{jt}^2 + \Pi_{12}\bar{P}^2\tilde{p}_{jt}E_{jt}\tilde{p}_{it} + \Pi_{13}\bar{P}^2\tilde{p}_{jt}E_{jt}\tilde{p}_t + \Pi_{14}\bar{P}\bar{Q}\tilde{p}_{jt}E_{jt}\tilde{y}_t + \Pi_{15}\bar{P}\bar{\varphi}_{it}\tilde{p}_{jt}E_{jt}\tilde{\phi}_{it} \\ &+ \text{other terms}\end{aligned}$$

where Π_1 is the partial derivative of profit with respect to the first argument (P_{jt}) and the $\Pi_{1\cdot}$ coefficients are the partial derivatives of Π_1, all evaluated at the perfect-information equilibrium values, and $\bar{P}$ is the perfect-information equilibrium price. The "other terms" are all other terms in the second-order approximation that do not depend on $\tilde{p}_{jt}$, which are irrelevent here since they will not affect the firm's pricing decision.

The problem faced by firm j can now be written

$$\max_{P_{jt}} \tilde{\Pi}_{jt}$$

And the solution is characterized by the first-order condition

$$\frac{\partial \tilde{\Pi}}{\partial \tilde{p}_{jt}} = 0 = \Pi_1 \bar{P} + \Pi_{11} \bar{P}^2 \tilde{p}_{jt} + \Pi_{12} \bar{P}^2 E_{jt} \tilde{p}_{it} + \Pi_{13} \bar{P}^2 E_{jt} \tilde{p}_t + \Pi_{14} \bar{P} \bar{Q} E_{jt} \tilde{Q}_t + \Pi_{15} \bar{P} \bar{\varphi}_{it} \tilde{E}_{jt} \tilde{\phi}_{it}$$

which reduces to

$$\begin{aligned} \tilde{p}^*_{jt} &= -\gamma E_{jt} \tilde{\phi}_{it} + \zeta E_{jt} \tilde{q}_t + (1 - \zeta) E_{jt} \tilde{p}_t \\ &= E_{jt} \tilde{p}^{\diamond}_{jt} \end{aligned}$$

Appendix C: Log Approximations

8.5 C-1: Log-linear approximation to the aggregate price index

For results that follow, we will require a log-linear approximation to the price index P_t. Recall from *A-1: Constant Elasticity of Substitution Preferences* that P_t is derived as the (minimum) cost of purchasing one unit of the consumption good and is defined to be

$$P_t = \left[\sum_{i=1}^{I} \mu_i P_{it}^{\frac{p}{p-1}} \right]^{\frac{p-1}{p}}$$

We take the log-linear approximation around the where all prices are the same $P_{jt} = P_{it} = P_t \equiv \bar{P}$. Define $\tilde{p}_t \equiv \frac{(P_t - \bar{P})}{\bar{P}} \approx \log\left(\frac{P_t}{\bar{P}}\right) = p_t - \bar{p}$ so that $\tilde{p}_t$ is the aggregate price in proportional deviation-from-steady-state form.

$$
\begin{aligned}
\bar{P} + (1)(P_t - \bar{P}) &= \bar{P} + \sum_{i=1}^{I} \frac{p-1}{p} P_t^{\frac{1}{1-p}} \mu_i P_{it}^{\frac{1}{p-1}} \frac{p}{p-1} (P_{it} - \bar{P}) \\
(P_t - \bar{P}) &= \sum_{i=1}^{I} \mu_i (P_{it} - \bar{P}) \\
\bar{P}\tilde{p}_t &= \sum_{i=1}^{I} \mu_i \bar{P}_i \tilde{p}_{it}
\end{aligned}
$$

Thus the log-approximation aggregate price is described by

$$
\tilde{p}_t = \sum_{i=1}^{I} \mu_i \tilde{p}_{it}
$$

8.6 C-2: Log-quadratic approximation to an intermediate good firm's profit function

Recall from *B-2: Optimal Price Setting* that the problem faced by firm j can be written

$$
\max_{P_{jt}} \Pi_{jt}(P_{jt}, P_{it}, P_t, Y_t, \varphi_{it})
$$

A second-order approximation to this objective function around the perfect information non-stochastic equilibrium is given by

$$
\begin{aligned}
\bar{\Pi} + (1)(\Pi_{jt} - \bar{\Pi}) =& \bar{\Pi} + \Pi_1 (P_{jt} - \bar{P}) + \frac{\Pi_{11}}{2!}(P_{jt} - \bar{P})^2 + \Pi_{12}(P_{jt} - \bar{P})(P_{it} - \bar{P}) \\
& + \Pi_{13}(P_{jt} - \bar{P}) E_{jt}(P_t - \bar{P}) + \Pi_{14}(P_{jt} - \bar{P}) E_{jt}(Y_t - \bar{Y}) + \Pi_{15}(P_{jt} - \bar{P}) E_{jt}(\varphi_{it} - \bar{\varphi}_t) \\
& + \text{other terms}
\end{aligned}
$$

where $\bar{P}$ denotes the price at which $P_{jt} = P_{it} = P_t \equiv \bar{P}$ and $\bar{Q}_t$ and $\bar{\varphi}_t$ denote the means of the processes. Π_1 is the partial derivative of profit with respect to the first argument Π_{1*} represent second partial derivatives, all evaluated at the point at which all prices are the same. The term "other terms" collects all terms that do not depend on $\tilde{p}_{jt}$ (irrelevent for our purposes since they do not affect the firm's pricing decision). The forms of these partial derivatives are derived below.

In log-deviation form the objective function is

$$\begin{aligned}\tilde{\Pi}(\tilde{p}_{jt}, \tilde{p}_{it}, \tilde{p}_t, \tilde{y}_t, \tilde{\phi}_{it}) =& \Pi_1 \bar{P}\tilde{p}_{jt} + \frac{\Pi_{11}}{2!}\bar{P}^2\tilde{p}_{jt}^2 + \Pi_{12}\bar{P}^2\tilde{p}_{jt}\tilde{p}_{it} + \Pi_{13}\bar{P}^2\tilde{p}_{jt}E_{jt}\tilde{p}_t + \Pi_{14}\bar{P}\bar{Y}\tilde{p}_{jt}E_{jt}\tilde{y}_t + \Pi_{15}\bar{P}\bar{\varphi}_t\tilde{p}_{jt}E_{jt}\tilde{\phi}\\ &+ \text{other terms}\end{aligned}$$

C-2.2: Derivatives

Below we calculate the first and second partial derivatives used in the log-quadratic approximation, above. The second partial derivatives are all first with respect to P_{jt} and then second with respect to the given variable. Evaluation of derivatives will be around the perfect information non-stochastic equilibrium (see *E-1: Perfect Information*) in which all prices are the same (we will use that the means of the shocks have been defined to be identical, see *Stochastic processes*, to guarantee the last condition).

Before calculating the derivatives, simplify the objective function by applying market clearing and household optimization so that it can be written

$$\begin{aligned}
\Pi_{jt} &= E_{jt}\left[U'(Y_t)\left(P_{jt} - \frac{W_{jt}}{\varphi_{it}}\right) Y_{jt}\right] \\
&= E_{jt}\left[Y_t^{-\sigma}\left(P_{jt} - \frac{W_{jt}}{\varphi_{it}}\right)\left(\frac{1}{\mu_i}\right)\left(\frac{P_{jt}}{P_{it}}\right)^{\frac{1}{r-1}} \mu_i \left(\frac{P_{it}}{P_t}\right)^{\frac{1}{p-1}} Y_t\right] \\
&= E_{jt}\left[\left(P_{jt} - \frac{W_{jt}}{\varphi_{it}}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma}\right] \\
&= E_{jt}\left[P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma}\left(P_{jt}^{1+\frac{1}{r-1}} - P_{jt}^{\frac{1}{r-1}}\frac{W_{jt}}{\varphi_{it}}\right)\right]
\end{aligned}$$

Since factor markets are perfectly competitive, so that firms are wage-takers, the wage cannot yet be substituted out.

First derivative, with respect to P_{jt}

$$\frac{\partial \Pi_{jt}}{\partial P_{jt}} = E_{jt}\left[\left(\frac{1}{r-1}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma} - \left(1 + \frac{1}{r-1}\right) P_{jt}^{\frac{1}{r-1}-1} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma}\frac{W_{jt}}{\varphi_{it}}\right]$$

After taking the derivative, wages can be substituted out using the firm's static first-order condition

$$\begin{aligned}
\frac{\partial \Pi_{jt}}{\partial P_{jt}} &= E_{jt}\left(\frac{r}{r-1}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma} - E_{jt}\left(\frac{1}{r-1}\right)\frac{1}{\varphi_{it}} P_{jt}^{\frac{1}{r-1}-1} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma}\left(P_t \frac{n_{jt}^{\varepsilon}}{Y_t^{-\sigma}}\right) \\
&= E_{jt}\left(\frac{r}{r-1}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma} - E_{jt}\left(\frac{1}{r-1}\right)\frac{1}{\varphi_{it}^{1+\varepsilon}} P_{jt}^{\frac{1}{r-1}-1} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{2-p}{1-p}} Y_t \left[Y_{jt}\right]^{\varepsilon} \\
&= E_{jt}\left(\frac{r}{r-1}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma} - E_{jt}\left(\frac{1}{r-1}\right)\frac{1}{\varphi_{it}^{1+\varepsilon}} P_{jt}^{\frac{1}{r-1}-1} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{2-p}{1-p}} Y_t \left[\left(\frac{1}{\mu_i}\right)\left(\frac{P_{jt}}{P_{it}}\right)\right.
\end{aligned}$$

This finally yields the first derivative of the log-quadratic approximation to the profit function with respect to price

$$\frac{\partial \Pi_{jt}}{\partial P_{jt}} = E_{jt}\left(\frac{r}{r-1}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma} - E_{jt}\left(\frac{1}{r-1}\right)\frac{1}{\varphi_{it}^{1+\varepsilon}} P_{jt}^{\frac{1+\varepsilon}{r-1}-1} P_{it}^{(1+\varepsilon)\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{2-p+\varepsilon}{1-p}} Y_t^{1+\varepsilon}$$

Evaluation at the point where all prices are the same implies $P_{jt} = P_{it} = P_t \equiv \bar{P}$. At this point we also have $\bar{\varphi}_{it} \equiv \bar{\varphi}_t$ for $i = 1, \ldots, I$. Then the exponent on the price on the left hand side is

$$\begin{aligned}
\frac{1}{r-1} + \frac{r-p}{(p-1)(r-1)} - \frac{1}{p-1} &= \frac{(p-1)+(r-p)-(r-1)}{(p-1)(r-1)} \\
&= 0
\end{aligned}$$

and the exponent on price on the right hand side is:

$$\begin{aligned}
\frac{1+\varepsilon}{r-1} - 1 + (1+\varepsilon)\frac{r-p}{(p-1)(r-1)} - \frac{1+\varepsilon}{p-1} + 1 &= (1+\varepsilon)\frac{(p-1)+(r-p)-(r-1)}{(p-1)(r-1)} - 1 + 1 \\
&= 0
\end{aligned}$$

this leads to

$$\Pi_1 \equiv \left.\frac{\partial \Pi_{jt}}{\partial P_{jt}}\right|_{\bar{P},\bar{Y},\bar{\varphi}_t} = \left(\frac{r}{r-1}\right)\bar{Y}^{1-\sigma} - \left(\frac{1}{r-1}\right)\left(\frac{1}{\bar{\varphi}_t}\right)^{1+\varepsilon}\bar{Y}^{1+\varepsilon}$$

Recall that $\bar{Y} = Y_t^n = r^{\frac{\alpha}{\zeta}}\left(\frac{1}{\bar{\varphi}_t}\right)^{-\frac{\gamma}{\zeta}}$ when shocks have a common mean (see *E-1: Perfect Information*). Then

$$\begin{aligned}
\Pi_1 \equiv \left.\frac{\partial \Pi_{jt}}{\partial P_{jt}}\right|_{\bar{P},\bar{Y},\bar{\varphi}_t} &= \left(\frac{r}{r-1}\right) r^{\frac{1-\sigma}{\sigma+\varepsilon}}\left(\frac{1}{\bar{\varphi}_t}\right)^{-(1-\sigma)\frac{1+\varepsilon}{\sigma+\varepsilon}} - \left(\frac{1}{r-1}\right)\left(\frac{1}{\bar{\varphi}_t}\right)^{1+\varepsilon}\left(\frac{1}{\bar{\varphi}_t}\right)^{-(1+\varepsilon)\frac{1+\varepsilon}{\sigma+\varepsilon}} r^{\frac{1+\varepsilon}{\sigma+\varepsilon}} \\
&= \left(\frac{1}{r-1}\right) r^{\frac{1-\sigma}{\sigma+\varepsilon}+1}\left(\frac{1}{\bar{\varphi}_t}\right)^{-(1-\sigma)\frac{1+\varepsilon}{\sigma+\varepsilon}} - \left(\frac{1}{r-1}\right)\left(\frac{1}{\bar{\varphi}_t}\right)^{1+\varepsilon-(1+\varepsilon)\frac{1+\varepsilon}{\sigma+\varepsilon}} r^{\frac{1+\varepsilon}{\sigma+\varepsilon}} \\
&= \left(\frac{1}{r-1}\right) r^{\frac{1+\varepsilon}{\sigma+\varepsilon}}\left(\frac{1}{\bar{\varphi}_t}\right)^{-(1-\sigma)\frac{1+\varepsilon}{\sigma+\varepsilon}} - \left(\frac{1}{r-1}\right)\left(\frac{1}{\bar{\varphi}_t}\right)^{1+\varepsilon-(1+\varepsilon)r^{\frac{1+\varepsilon}{\sigma+\varepsilon}}\frac{1+\varepsilon}{\sigma+\varepsilon}} \\
&= 0
\end{aligned}$$

where the last equality can be found by noting that the right-hand side exponent on $1/\varphi$ is the same as the left-hand side exponent:

$$(1+\varepsilon) - (1+\varepsilon)\frac{1+\varepsilon}{\sigma+\varepsilon} = (1+\varepsilon)\left[1 - \frac{1+\varepsilon}{\sigma+\varepsilon}\right] = (1+\varepsilon)\left[\frac{\sigma+\varepsilon-1-\varepsilon}{\sigma+\varepsilon}\right] = (1+\varepsilon)\left[\frac{\sigma-1}{\sigma+\varepsilon}\right]$$
$$= -(1+\varepsilon)\left[\frac{1-\sigma}{\sigma+\varepsilon}\right] = -(1-\sigma)\left[\frac{1+\varepsilon}{\sigma+\varepsilon}\right]$$

Second derivatives

Note that

$$\begin{aligned}\bar{W} &= \frac{1}{\bar{\varphi}_t^{\varepsilon}}\bar{P}\bar{Y}^{\sigma}\left[\left(\frac{1}{\mu_i}\right)\left(\frac{\bar{P}}{\bar{\bar{P}}}\right)^{\frac{1}{r-1}}\mu_i\left(\frac{\bar{P}}{\bar{\bar{P}}}\right)^{\frac{1}{p-1}}\bar{Y}\right]^{\varepsilon} \\ &= \frac{1}{\bar{\varphi}_t^{\varepsilon}}\bar{P}\bar{Y}^{\sigma+\varepsilon} \\ &= r\bar{\varphi}_t\bar{P}\end{aligned}$$

Second derivative, with respect to P_{jt}

$$\frac{\partial^2\Pi_{jt}}{\partial P_{jt}^2} = \frac{\partial}{\partial P_{jt}}E_{jt}\left(\frac{r}{r-1}\right)P_{jt}^{\frac{1}{r-1}}P_{it}^{\frac{r-p}{(p-1)(r-1)}}P_t^{\frac{1}{1-p}}Y_t^{1-\sigma} - \left(\frac{1}{r-1}\right)P_{jt}^{\frac{1}{r-1}-1}P_{it}^{\frac{r-p}{(p-1)(r-1)}}P_t^{\frac{1}{1-p}}Y_t^{1+\varepsilon}\frac{1}{\varphi_{it}^{1+\varepsilon}}P_t\left[\left(\frac{P_{jt}}{P_{it}}\right)\right.$$
$$= E_{jt}\left(\frac{r}{r-1}\right)\left(\frac{1}{r-1}\right)P_{jt}^{\frac{1+\varepsilon}{r-1}-1}P_{it}^{\frac{r-p}{(p-1)(r-1)}}P_t^{\frac{1}{1-p}}Y_t^{1-\sigma} - \left(\frac{1}{r-1}\right)\left(\frac{1}{r-1}-1\right)P_{jt}^{\frac{1}{r-1}-2}P_{it}^{\frac{r-p}{(p-1)(r-1)}}P_t^{\frac{1}{1-p}}$$

$$
\begin{aligned}
\Pi_{11} \equiv \left.\frac{\partial^2 \Pi_{jt}}{\partial P_{jt}^2}\right|_{\bar{P},\bar{Y},\bar{\varphi}_t} &= \left(\frac{r}{r-1}\right)\left(\frac{1}{r-1}\right)\bar{P}^{-1}\bar{Y}^{1-\sigma} - \left(\frac{1}{r-1}\right)\left(\frac{2-r+\varepsilon}{r-1}\right)\bar{P}^{-2}\bar{Y}^{1-\sigma}\frac{1}{\bar{\varphi}_t}\bar{W} \\
&= \left(\frac{1}{r-1}\right)^2 \bar{P}^{-1}\bar{Y}^{1-\sigma}\left[r - (2-r+\varepsilon)\frac{1}{\bar{\varphi}_t^{1+\varepsilon}}\bar{Y}^{\sigma+\varepsilon}\right] \\
&= \left(\frac{1}{r-1}\right)^2 \bar{P}^{-1}\bar{Y}^{1-\sigma} r\left[1 - (2-r+\varepsilon)\right] \\
&= \left(\frac{1}{r-1}\right)^2 \bar{P}^{-1}\bar{Y}^{1-\sigma} r\left[r - 1 - \varepsilon\right]
\end{aligned}
$$

Notice that since $r \in [0,1)$, then this term is strictly negative.

Second derivative, with respect to P_{it}

$$
\begin{aligned}
\frac{\partial^2 \Pi_{jt}}{\partial P_{jt}\partial P_{it}} &= \frac{\partial}{\partial P_{it}} E_{jt}\left(\frac{r}{r-1}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma} - \left(\frac{1}{r-1}\right) P_{jt}^{\frac{1}{r-1}-1} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1+\varepsilon} \frac{1}{\varphi_{it}^{1+\varepsilon}} P_t \Bigg[\Bigg(\\
&= E_{jt}\left(\frac{r}{r-1}\right)\left(\frac{r-p}{(p-1)(r-1)}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}-1} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma} - \left(\frac{1}{r-1}\right)\left(\frac{(1+\varepsilon)(r-p)}{(p-1)(r-1)}\right) P_{jt}^{\frac{1}{r-1}}
\end{aligned}
$$

$$
\begin{aligned}
\Pi_{12} \equiv \left.\frac{\partial^2 \Pi_{jt}}{\partial P_{jt}\partial P_{it}}\right|_{\bar{P},\bar{Y},\bar{\varphi}_t} &= \left(\frac{r}{r-1}\right)\left(\frac{r-p}{(p-1)(r-1)}\right)\bar{P}^{-1}\bar{Y}^{1-\sigma} - \left(\frac{1}{r-1}\right)\left(\frac{(1+\varepsilon)(r-p)}{(p-1)(r-1)}\right)\bar{P}^{-2}\bar{Y}^{1-\sigma}\frac{\bar{W}}{\bar{\varphi}_t} \\
&= \frac{r-p}{(p-1)(r-1)^2}\bar{P}^{-1}\bar{Y}^{1-\sigma}\left[r - (1+\varepsilon)\frac{1}{\bar{\varphi}_t^{1+\varepsilon}}\bar{Y}^{\sigma+\varepsilon}\right] \\
&= \frac{r-p}{(p-1)(r-1)^2}\bar{P}^{-1}\bar{Y}^{1-\sigma} r\left[1 - (1+\varepsilon)\right] \\
&= \frac{r-p}{(p-1)(r-1)^2}\bar{P}^{-1}\bar{Y}^{1-\sigma} r\left[-\varepsilon\right]
\end{aligned}
$$

Second derivative, with respect to P_t

$$\frac{\partial^2 \Pi_{jt}}{\partial P_{jt} \partial P_t} = \frac{\partial}{\partial P_t} E_{jt} \left(\frac{r}{r-1}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma} - \left(\frac{1}{r-1}\right) P_{jt}^{\frac{1}{r-1}-1} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1+\varepsilon} \frac{1}{\varphi_{it}^{1+\varepsilon}} P_t \left[\left(\frac{P_{jt}}{P_{it}} \right.\right.$$

$$= E_{jt} \left(\frac{r}{r-1}\right) \left(\frac{1}{1-p}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}-1} Y_t^{1-\sigma} - \left(\frac{1}{r-1}\right) \left(\frac{1}{1-p} + 1 + \frac{\varepsilon}{1-p}\right) P_{jt}^{\frac{1}{r-1}-1} P_{it}^{\frac{}{(p-}}$$

$$\begin{aligned}
\Pi_{13} \equiv \left. \frac{\partial^2 \Pi_{jt}}{\partial P_{jt} \partial P_t} \right|_{\bar{P}, \bar{Y}, \bar{\varphi}_t} &= \left(\frac{r}{r-1}\right) \left(\frac{1}{1-p}\right) \bar{P}^{-1} \bar{Y}^{1-\sigma} - \left(\frac{1}{r-1}\right) \left(\frac{2-p+\varepsilon}{1-p}\right) \bar{P}^{-1} \bar{Y}^{1+\varepsilon} \frac{\bar{W}}{\bar{P} \bar{\varphi}_t} \\
&= \frac{1}{(1-p)(r-1)} \bar{P}^{-1} \bar{Y}^{1-\sigma} \left[r - (2-p+\varepsilon) \frac{1}{\bar{\varphi}_t^{1+\varepsilon}} \bar{Y}^{\sigma+\varepsilon} \right] \\
&= -\frac{1}{(p-1)(r-1)} \bar{P}^{-1} \bar{Y}^{1-\sigma} r \left[-1 + p - \varepsilon \right] \\
&= (-1) \alpha^{-1} \frac{1}{r-1} \bar{P}^{-1} \bar{Y}^{1-\sigma} r
\end{aligned}$$

Second derivative, with respect to Y_t

$$\frac{\partial^2 \Pi_{jt}}{\partial P_{jt} \partial Y_t} = \frac{\partial}{\partial Y_t} E_{jt} \left(\frac{r}{r-1}\right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma} - \left(\frac{1}{r-1}\right) P_{jt}^{\frac{1}{r-1}-1} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1+\varepsilon} \frac{1}{\varphi_{it}^{1+\varepsilon}} P_t \left[\left(\frac{P_{jt}}{P_{it}} \right.\right.$$

$$= E_{jt} \left(\frac{r}{r-1}\right) (1-\sigma) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{-\sigma} - \left(\frac{1}{r-1}\right) (1+\varepsilon) P_{jt}^{\frac{1}{r-1}-1} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{\varepsilon} \frac{1}{\varphi_{it}^{1+\varepsilon}} P_t$$

$$
\begin{aligned}
\Pi_{14} \equiv \left. \frac{\partial^2 \Pi_{jt}}{\partial P_{jt} \partial Y_t} \right|_{\bar{P}, \bar{Y}, \bar{\varphi}_t} &= \left(\frac{r}{r-1} \right) (1-\sigma) \bar{Y}^{-\sigma} - \left(\frac{1}{r-1} \right) (1+\varepsilon) \bar{P}^{-1} \bar{Y}^{\varepsilon} \frac{1}{\bar{\varphi}_t^{1+\varepsilon}} \\
&= \frac{1}{r-1} \bar{Y}^{-\sigma} \left[(1-\sigma) r - (1+\varepsilon) \frac{1}{\bar{\varphi}_t^{1+\varepsilon}} \bar{Y}^{\sigma+\varepsilon} \right] \\
&= \frac{1}{r-1} \bar{Y}^{-\sigma} r \left[(1-\sigma) - (1+\varepsilon) \right] \\
&= (-1) \frac{1}{r-1} \bar{Y}^{-\sigma} r \left[\varepsilon + \sigma \right]
\end{aligned}
$$

Second derivative, with respect to φ_t

$$
\begin{aligned}
\frac{\partial^2 \Pi_{jt}}{\partial P_{jt} \partial \varphi_t} &= \frac{\partial}{\partial \varphi_{it}} E_{jt} \left(\frac{r}{r-1} \right) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1-\sigma} - \left(\frac{1}{r-1} \right) P_{jt}^{\frac{1}{r-1}-1} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1+\varepsilon} \frac{1}{\varphi_{it}^{1+\varepsilon}} P_t \left[\left(\frac{P_{jt}}{P_{it}} \right. \right. \\
&= \left(\frac{1}{r-1} \right) (1+\varepsilon) P_{jt}^{\frac{1}{r-1}} P_{it}^{\frac{r-p}{(p-1)(r-1)}} P_t^{\frac{1}{1-p}} Y_t^{1+\varepsilon} \frac{1}{\varphi^{2+\varepsilon}} \left[\left(\frac{P_{jt}}{P_{it}} \right)^{\frac{1}{r-1}} \left(\frac{P_{it}}{P_t} \right)^{\frac{1}{p-1}} \right]^{\varepsilon}
\end{aligned}
$$

$$
\begin{aligned}
\Pi_{15} \equiv \left. \frac{\partial^2 \Pi_{jt}}{\partial P_{jt} \partial \varphi_t} \right|_{\bar{P}, \bar{Y}, \bar{\varphi}_t} &= \left(\frac{1}{r-1} \right) (1+\varepsilon) \bar{Y}^{1+\varepsilon+\sigma-\sigma} \frac{1}{\bar{\varphi}_t^{2+\varepsilon}} \\
&= \left(\frac{1+\varepsilon}{r-1} \right) r \bar{Y}^{1-\sigma} \frac{1}{\bar{\varphi}_t}
\end{aligned}
$$

8.7 C-3: Log-linear approximation to an intermediate good firm's pricing decision

Taking first-order conditions with respect to the firm's objective function (8.6) yields

$$
\begin{aligned}
\frac{\partial \tilde{\Pi}}{\partial \tilde{p}_{jt}} = 0 =& \Pi_1 \bar{P} + \Pi_{11} \bar{P}^2 \tilde{p}_{jt} + \Pi_{12} \bar{P}^2 \tilde{p}_{it} + \Pi_{13} \bar{P}^2 E_{jt} \tilde{p}_t + \Pi_{14} \bar{P} \bar{Y} E_{jt} \tilde{y}_t + \Pi_{15} \bar{P} \bar{\varphi}_t \tilde{E}_{jt} \tilde{\phi}_{it} \\
=& 0 \\
& + \Pi_{11} \bar{P}^2 \tilde{p}_{jt} \\
& + \Pi_{12} \bar{P}^2 \tilde{p}_{it} \\
& + \Pi_{13} \bar{P}^2 E_{jt} \tilde{p}_t \\
& + \Pi_{14} \bar{P} \bar{Y} E_{jt} \tilde{y}_t \\
& + \Pi_{15} \bar{P} \bar{\varphi}_t E_{jt} \tilde{\phi}_{it} \\
=& 0 \\
& + \left(\frac{1}{r-1} \right)^2 \bar{P}^{-1} \bar{Y}^{1-\sigma} r \left[r - 1 - \varepsilon \right] \bar{P}^2 \tilde{p}_{jt} \\
& + \frac{r-p}{(p-1)(r-1)^2} \bar{P}^{-1} \bar{Y}^{1-\sigma} r \left[-\varepsilon \right] \bar{P}^2 \tilde{p}_{it} \\
& + (-1) \alpha^{-1} \frac{1}{r-1} \bar{P}^{-1} \bar{Y}^{1-\sigma} r \bar{P}^2 E_{jt} \tilde{p}_t \\
& + (-1) \frac{1}{r-1} \bar{Y}^{-\sigma} r \left[\varepsilon + \sigma \right] \bar{P} \bar{Y} E_{jt} \tilde{y}_t \\
& + \left(\frac{1+\varepsilon}{r-1} \right) r \bar{Y}^{1-\sigma} \frac{1}{\bar{\varphi}_t} \bar{P} \bar{\varphi}_t E_{jt} \tilde{\phi}_{it} \\
=& + \left(\frac{1}{r-1} \right)^2 \left[r - 1 - \varepsilon \right] \tilde{p}_{jt} \\
& + \frac{r-p}{(p-1)(r-1)^2} \left[-\varepsilon \right] \tilde{p}_{it} \\
& + (-1) \alpha^{-1} \frac{1}{r-1} E_{jt} \tilde{p}_t \\
& + (-1) \frac{1}{r-1} \left[\varepsilon + \sigma \right] E_{jt} \tilde{y}_t \\
& + \left(\frac{1+\varepsilon}{r-1} \right) E_{jt} \tilde{\phi}_{it}
\end{aligned}
$$

Notice that $P_{it} = P_{jt}$ since all firms within an industry face the same problem, and also that:

$$
\begin{aligned}
(p-1)(r-1-\varepsilon)+(r-p)(-\varepsilon) &= pr - p - p\varepsilon - r + 1 + \varepsilon - r\varepsilon + p\varepsilon \\
&= r(p-1-\varepsilon) - (p-1-\varepsilon) \\
&= (r-1)(p-1-\varepsilon)
\end{aligned}
$$

and $\frac{p-1-\varepsilon}{p-1} = \frac{1-p+\varepsilon}{1-p} \equiv \alpha^{-1}$. Then we have

$$
\begin{aligned}
0 =& \frac{1}{(p-1)(r-1)^2}\left[(p-1)(r-1-\varepsilon)+(r-p)(-\varepsilon)\right]\tilde{p}_{jt} \\
&+ (-1)\alpha^{-1}\frac{1}{r-1}E_{jt}\tilde{p}_t \\
&+ (-1)\frac{1}{r-1}\left[\varepsilon+\sigma\right]E_{jt}\tilde{y}_t \\
&+ \left(\frac{1+\varepsilon}{r-1}\right)E_{jt}\tilde{\phi}_{it} \\
=& \alpha^{-1}\tilde{p}_{jt} \\
&+ (-1)\alpha^{-1}E_{jt}\tilde{p}_t \\
&+ (-1)(\varepsilon+\sigma)E_{jt}\tilde{y}_t \\
&+ (1+\varepsilon)E_{jt}\tilde{\phi}_{it}
\end{aligned}
$$

and finally this reduces to the firms' imperfect-information pricing rule

$$
\begin{aligned}
\tilde{p}^*_{jt} &= -\gamma E_{jt}\tilde{\phi}_{it} + \zeta E_{jt}\tilde{y}_t + E_{jt}\tilde{p}_t \\
&= -\gamma E_{jt}\tilde{\phi}_{it} + \zeta E_{jt}\tilde{q}_t + (1-\zeta)E_{jt}\tilde{p}_t \\
&= E_{jt}p^{\diamond}_{jt}
\end{aligned}
$$

8.8 C-4: Log-quadatic approximation to profit loss due to imperfect information

We cconsider the loss in profit from a firm setting a non-profit-maximizing price $p^*_{jt} = E_{jt}p^{\diamond}_{jt}$. The log-quadratic approximation to the profit function is

$$
\begin{aligned}
\tilde{\Pi}(\tilde{p}_{jt}, \tilde{p}_{it}, \tilde{p}_t, \tilde{y}_t, \tilde{\phi}_{it}) =& \Pi_1\bar{P}\tilde{p}_{jt} + \frac{\Pi_{11}}{2!}\bar{P}^2\tilde{p}^2_{jt} + \Pi_{12}\bar{P}^2\tilde{p}_{jt}\tilde{p}_{it} + \Pi_{13}\bar{P}^2\tilde{p}_{jt}E_{jt}\tilde{p}_t + \Pi_{14}\bar{P}\bar{Y}\tilde{p}_{jt}E_{jt}\tilde{y}_t + \Pi_{15}\bar{P}\bar{\varphi}_t\tilde{p}_{jt}E_{jt}\tilde{\phi} \\
&+ \text{other terms}
\end{aligned}
$$

and recall that the "other terms" do not depend on the firm's price decision. The loss in profits is then

$$
\begin{aligned}
\tilde{\Pi}(\tilde{p}^{\diamond}_{jt}, \cdot) - \tilde{\Pi}(\tilde{p}^*_{jt}, \cdot) =& \Pi_1\bar{P}(\tilde{p}^{\diamond}_{jt} - \tilde{p}^*_{jt}) \\
&+ \left(\frac{\Pi_{11}}{2}\right)\bar{P}^2\left(\tilde{p}^{\diamond^2}_{jt} - \tilde{p}^{*^2}_{jt}\right) \\
&+ (\Pi_{12}\bar{P}^2\tilde{p}_{it} + \Pi_{13}\bar{P}^2\tilde{p}_t + \Pi_{14}\bar{P}\bar{Y}\tilde{Y}_t + \Pi_{15}\bar{P}\bar{\varphi}_t\tilde{\phi}_{it})(\tilde{p}^{\diamond}_{jt} - \tilde{p}^*_{jt})
\end{aligned}
$$

Note first that $\Pi_1 = 0$, and second, from the first-order condition above in the perfect-information case, that $\Pi_{12}\bar{P}^2\tilde{p}_{it} + \Pi_{13}\bar{P}^2\tilde{p}_t + \Pi_{14}\bar{P}\bar{Y}\tilde{Y}_t + \Pi_{15}\bar{P}\bar{\varphi}_t\tilde{\phi}_t = -\Pi_{11}\bar{P}^2\tilde{p}^{\diamond}_{jt}$. Then we can rewrite the loss in profits as

$$
\begin{aligned}
\tilde{\Pi}(\tilde{p}^{\diamond}_{jt}, \cdot) - \tilde{\Pi}(\tilde{p}^*_{jt}, \cdot) &= \bar{P}^2\left(\frac{\Pi_{11}}{2}\right)\left(\tilde{p}^{\diamond^2}_{jt} - \tilde{p}^{*^2}_{jt}\right) - \bar{P}^2\Pi_{11}\tilde{p}^{\diamond}_{jt}(\tilde{p}^{\diamond}_{jt} - \tilde{p}^*_{jt}) \\
&= -\bar{P}^2\left(\frac{\Pi_{11}}{2}\right)\left(\tilde{p}^{\diamond^2}_{jt} + \tilde{p}^{*^2}_{jt}\right) + \bar{P}^2\Pi_{11}\tilde{p}^{\diamond}_{jt}\tilde{p}^*_{jt} \\
&= \left(-\frac{\Pi_{11}}{2}\bar{P}^2\right)\left(\tilde{p}^{\diamond}_{jt} - \tilde{p}^*_{jt}\right)^2
\end{aligned}
$$

Finally recall, from above, that $\Pi_{11} < 0$ so that the above is positive overall, indicating setting $\tilde{p}^*_{jt}$ yields less profits than setting $\tilde{p}^{\diamond}_{jt}$. The expected loss in profits due to imperfect information can be written

$$E_{jt}\left[\tilde{\Pi}(\tilde{p}^{\diamond}_{jt},\cdot)-\tilde{\Pi}(\tilde{p}^{*}_{jt},\cdot)\right]=-\hat{\Pi}_{11}\left(\tilde{p}^{\diamond}_{jt}-\tilde{p}^{*}_{jt}\right)^2$$

where $\hat{\Pi}_{11}=\left(\frac{\Pi_{11}}{2}\bar{P}^2\right)<0$.

Appendix D: Information Theory

This appendix collects information theoretic results.

8.9 D1: Mutual Information of Random Vectors

In the case that the variables are independent so that is no reduction in uncertainty, then $I(X;S)=H(X)-H(X)=0$. Supposing that $\mathbf{X}$ and $\mathbf{S}$ are finite n-dimensional independent vectors such that X_i and S_j are independent if $i\neq j$, then

$$\begin{aligned}
I(\mathbf{X};\mathbf{S}) &= H(X_1,\cdots,X_n)-H(X_1,\cdots,X_n|S_1,\cdots,S_n)\\
&=\sum_{i=1}^{n}H(X_i)+\sum_{i=1}^{n}H(X_i)-H(X_1,\cdots,X_n,S_1,\cdots,S_n)\\
&=\sum_{i=1}^{n}H(X_i)+\sum_{i=1}^{n}H(X_i)-\sum_{i=1}^{n}H(X_i,S_i)\\
&=\sum_{i=1}^{n}I(X_i;S_i)
\end{aligned}$$

where the third equality follows from an iterative application of the chain rule.

8.10 D2: Gaussian Mutual Information

Suppose, as we do above, that we have two mutually Gaussian random variables, ω and s^{ω} such that

$$s^{(\omega)} = \omega + \psi$$

where ψ is Gaussian white noise. A well-known result (see for example Cover and Thomas, 2006) for Gaussian random processes is that mutual information can be simply expressed.

$$\mathcal{I}(\omega, s^{(\omega)}) = \frac{1}{2}\log_2\left(\frac{1}{1-\rho^2_{\omega s^{(\omega)}}}\right)$$

where $\rho^2_{\omega s^{(\omega)}}$ is the correlation coefficent between the two processes. Now, notice that the correlation coefficient can be rewritten in terms of the processes variances

$$\begin{aligned}\rho^2_{\omega s^{(\omega)}} &= \left[\frac{Cov(\omega, s^{(\omega)})}{\sigma_\omega \sigma_{s^{(\omega)}}}\right]^2 = \left[\frac{\sigma^2_\omega}{\sigma_\omega \sigma_{s^{(\omega)}}}\right]^2 = \frac{\sigma^2_\omega}{\sigma^2_{s^{(\omega)}}} \\ &= \frac{\sigma^2_\omega}{\sigma^2_\omega + \sigma^2_\psi}\end{aligned}$$

and $1-\rho^2_{\omega s^{(\omega)}} = \frac{\sigma_\psi}{\sigma^2_\omega+\sigma^2_\psi}$. Then the mutual information can be rewritten also in terms of the processes variances

$$\begin{aligned}\mathcal{I}(\omega, s^{(\omega)}) &= \frac{1}{2}\log_2\left(\frac{\sigma^2_\omega + \sigma^2_\psi}{\sigma^2_\psi}\right) \\ &= \frac{1}{2}\log_2\left(\frac{\sigma^2_\omega}{\sigma^2_\psi} + 1\right)\end{aligned}$$

8.11 D3: Expressions for Mutual Information

Define the mutual information as $\kappa^{(\omega)}_j \equiv \mathcal{I}(\omega, s^{(\omega)}_j)$. Assume the processes are defined as above.

1. Using this, we can find an expression for the ratio of the variances in terms as a function of a level of mutual information:

$$2^{2\kappa_j^{(\omega)}} - 1 = \frac{\sigma_\omega^2}{\sigma_\psi^2}$$

2. Immediately we also have an expression for the variance of the signal in terms of a level of mutual information and the variance of the fundamental

$$\sigma_\psi^2 = \left(2^{2\kappa_j^{(\omega)}} - 1\right)^{-1} \sigma_\omega^2$$

3. The ratio of the variance of the fundamental to the variance of the signal, a key term in typical signal extraction results, can be derived from (2)

$$\begin{aligned} 2^{2\kappa_j^{(\omega)}} \sigma_\psi^2 &= \sigma_\omega^2 + \sigma_\psi^2 \\ 2^{-2\kappa_j^{(\omega)}} &= \frac{\sigma_\psi^2}{\sigma_\omega^2 + \sigma_\psi^2} \\ 1 - 2^{-2\kappa_j^{(\omega)}} &= \frac{\sigma_\omega^2}{\sigma_\omega^2 + \sigma_\psi^2} \end{aligned}$$

4. Finally, from (1) - (3) follows a result that will useful in expanding the expected loss of profits from setting a price with imperfect information

$$\left(1-2^{-2\kappa_j^{(\omega)}}\right)^2 = \left(\frac{2^{2\kappa_j^{(\omega)}}-1}{2^{2\kappa_j^{(\omega)}}}\right)^2$$
$$\left(1-2^{-2\kappa_j^{(\omega)}}\right)^2 \sigma_\psi^2 = \left(\frac{2^{2\kappa_j^{(\omega)}}-1}{2^{2\kappa_j^{(\omega)}}}\right)^2 \left(2^{2\kappa_j^{(\omega)}}-1\right)^{-1} \sigma_\omega^2$$
$$\left(1-2^{-2\kappa_j^{(\omega)}}\right)^2 \sigma_\psi^2 = \left(\frac{2^{2\kappa_j^{(\omega)}}-1}{2^{4\kappa_j^{(\omega)}}}\right) \sigma_\omega^2$$

Appendix E: Equilibrium

8.12 E-1: Perfect Information

In the perfect information case, firms set prices according to (see *B-2: Optimal Price Setting*):

$$p_{jt}^{\diamond} = \alpha \log \frac{1}{r} - \gamma \phi_{it} + \zeta q_t + (1-\zeta) p_t$$

Applying symmetry across firms within each industry, integrating across all firms, and using the log-linear approximation to the aggregate price index around the point where all prices are the same (see *C-1: Log-linear approximation to the aggregate price index*) yields

$$\int_J p_{jt} dj = \int_J \alpha \log \frac{1}{r} dj - \int_J \gamma \phi_{it} dj + \int_J \zeta y_t dj + \int_J p_t dj$$
$$\sum_{l=1}^{I} \mu_l p_{lt} = \alpha \log \frac{1}{r} - \gamma \sum_{l=1}^{I} \mu_l \phi_{lt} + \zeta y_t + p_t$$
$$p_t = \alpha \log \frac{1}{r} - \gamma \sum_{l=1}^{I} \mu_l \phi_{lt} + \zeta y_t + p_t$$

The perfect information equilibrium level of (real) output is then

$$y_t^{\diamond} = -\frac{\alpha}{\zeta}\log\frac{1}{r} + \frac{\gamma}{\zeta}\sum_{l=1}^{I}\mu_l\phi_{lt}$$

If nominal output is assumed to be an exogenous process $Q_t = P_tY_t$ then the perfect information equilibrium aggregate price level is

$$\begin{aligned} p_t^{\diamond} &= q_t - y_t^n \\ &= q_t + \frac{\alpha}{\zeta}\log\frac{1}{r} - \frac{\gamma}{\zeta}\sum_{l=1}^{I}\mu_l\phi_{lt} \end{aligned}$$

In deviation from steady-state form, these are written

$$\tilde{y}_t^{\diamond} = \frac{\gamma}{\zeta}\sum_{l=1}^{I}\mu_l\tilde{\phi}_{lt}$$

$$\tilde{p}_t^{\diamond} = \tilde{q}_t - \frac{\gamma}{\zeta}\sum_{l=1}^{I}\mu_l\tilde{\phi}_{lt}$$

And plugging this into the firms' pricing decision yields

$$\tilde{p}_{jt}^{\diamond} = -\gamma\tilde{\phi}_{it} + \zeta\tilde{q}_t + (1-\zeta)p_t^{\diamond}$$

Non-stochastic equilibrium

In the non-stochastic case where shocks are set to their means, we have

$$y_t^n = -\frac{\alpha}{\zeta}\log\frac{1}{r} + \frac{\gamma}{\zeta}\sum_{l=1}^{I}\mu_l\bar{\phi}_{lt}$$

$$p_t^n = \bar{q}_t + \frac{\alpha}{\zeta}\log\frac{1}{r} - \frac{\gamma}{\zeta}\sum_{l=1}^{I}\mu_l\bar{\phi}_{lt}$$

Plug in the aggregate price level to the firm's pricing rule

$$\begin{aligned}
p_{jt}^n &= \alpha\log\frac{1}{r} - \gamma\bar{\phi}_{it} + \zeta\bar{q}_t + (1-\zeta)\left[\bar{q}_t + \frac{\alpha}{\zeta}\log\frac{1}{r} - \frac{\gamma}{\zeta}\sum_{l=1}^{I}\mu_l\bar{\phi}_{lt}\right] \\
&= \left(1 + \frac{1-\zeta}{\zeta}\right)\alpha\log\frac{1}{r} - \gamma\bar{\phi}_{it} + (\zeta + 1 - \zeta)\bar{q}_t + (1-\zeta)\frac{\gamma}{\zeta}\sum_{l=1}^{I}\mu_l\bar{\phi}_{lt} \\
&= \bar{q}_t + \frac{\alpha}{\zeta}\log\frac{1}{r} - (1-\zeta)\frac{\gamma}{\zeta}\sum_{l=1}^{I}\mu_l\bar{\phi}_{lt} - \gamma\bar{\phi}_{it}
\end{aligned}$$

which yields the non-stochastic equilibrium price rule

$$p_{jt}^n = p_t^n + \gamma\left[\sum_{l=1}^{I}\mu_l\bar{\phi}_{lt} - \bar{\phi}_{it}\right]$$

To confirm that this is an equilibrium, integrate this pricing rule over all firms and use the log-linear approximation to the aggregate price index.

Notice that if all shocks are eliminated so that $\phi_{it} = 0$ for $i = 1, \ldots, I$ or if the shock is purely aggregate so that $\phi_{it} = \phi_{i't}$ for $i, i' = 1, \ldots, I$ then the perfect information equilibrium corresponds to point where all prices are the same and are equal to the aggregate price index (this last point is by construction, see *A-2: Budget Contraints*).

8.13 E-2: Rational Inattention under Gaussian White Noise

Here we follow a guess and verify approach. Given the form of the perfect information equilibrium, we guess that the equilibrium aggregate price level is given by

$$\tilde{p}_t^* = a\tilde{q}_t - \frac{\gamma}{\zeta}\sum_{l=1}^{I} b_l \mu_l \tilde{\phi}_{lt}$$

where a and $\{b_l\}_{l=1}^{I}$ are coefficients governing the extent of adjustment of the price level due to shocks.

E-2.1: Imperfect Information Pricing Rule

With this guess, firms' imperfect-information optimal price rule is

$$\begin{aligned}
\tilde{p}_{jt}^* &= E_{jt}\tilde{p}_{jt}^{\diamond} \\
&= -\gamma E_{jt}\tilde{\phi}_{it} + \zeta E_{jt} q_t + (1-\zeta) E_{jt}\tilde{p}_t \\
&= -\gamma E_{jt}\tilde{\phi}_{it} + \zeta E_{jt} q_t + (1-\zeta)\left[a E_{jt}\tilde{q}_t - \frac{\gamma}{\zeta}\sum_{l=1}^{I} b_l \mu_l E_{jt}\tilde{\phi}_{lt} \right] \\
&= \left[(1-\zeta)a + \zeta\right] E_{jt}\tilde{q}_t - (1-\zeta)\frac{\gamma}{\zeta}\sum_{l=1}^{I} b_l \mu_l \tilde{\phi}_{lt} - \gamma E_{jt}\tilde{\phi}_{it}
\end{aligned}$$

To ease notation in the following optimization problem in which the constant term simply be carried around, define $\bar{a} \equiv [(1-\zeta)a + \zeta]$ and $\bar{b}_l \equiv (1-\zeta)\frac{\gamma}{\zeta} b_l$; they will be unpacked again at the end to aid interpretation. Then the optimal price rule is

$$\tilde{p}_{jt}^* = \bar{a} E_{jt}\tilde{q}_t - \sum_{l=1}^{I} \bar{b}_l \mu_l \tilde{\phi}_{lt} - \gamma E_{jt}\tilde{\phi}_{it}$$

E-2.2: Price-Setting Mean Squared Error

Given the signals the firm receives, we can solve the expectations using typical signal extraction results

$$\tilde{p}^*_{jt} = \bar{a}\left(\frac{\sigma_q^2}{\sigma_q^2+\sigma^2_{\psi_j^{(q)}}}\right)s^q_{jt} - \sum_{l\neq i}\bar{b}_l\mu_l\left(\frac{\sigma_l^2}{\sigma_l^2+\sigma^2_{\psi_j^{(l)}}}\right)s^{(l)}_{jt} - \left(\bar{b}_i\mu_i+\gamma\right)\left(\frac{\sigma_i^2}{\sigma_i^2+\sigma^2_{\psi_j^{(i)}}}\right)s^{(i)}_{jt}$$

Using results from *D3: Expressions for Mutual Information*, we can rewrite this in terms of mutual information as

$$\begin{aligned}\tilde{p}^*_{jt} =& \bar{a}\left(1-2^{-2\kappa_j^{(q)}}\right)\left(\tilde{q}_t-\psi^{(q)}_{jt}\right)\\ &-\sum_{l\neq i}\bar{b}_l\mu_l\left(1-2^{-2\kappa_j^{(l)}}\right)\left(\tilde{\phi}_{lt}-\psi^{(l)}_{jt}\right)\\ &-\left(\bar{b}_i\mu_i+\gamma\right)\left(1-2^{-2\kappa_j^{(i)}}\right)\left(\tilde{\phi}_{it}-\psi^{(i)}_{jt}\right)\end{aligned}$$

The expected loss in profits from setting an imperfect-information price is

$$E_{jt}\left[\tilde{\Pi}_{jt}\left(\tilde{p}^{\diamond}_{jt},\cdot\right)-\tilde{\Pi}_{jt}\left(\tilde{p}^*_{jt},\cdot\right)\right] = \left(\frac{\hat{\Pi}_{11}}{2}\right)E_{jt}\left[\left(\tilde{p}^{\diamond}_{jt}-\tilde{p}^*_{jt}\right)^2\right]$$

which is a constant times the mean squared error of the imperfect-information price. The difference between the perfect- and imperfect-information prices is

$$\begin{aligned}
\tilde{p}^{\diamond}_{jt} - \tilde{p}^{*}_{jt} =& \bar{a}\left[\tilde{q}_t - \left(1 - 2^{-2\kappa_j^{(q)}}\right)\left(\tilde{q}_t + \psi^{(q)}_{jt}\right)\right] \\
& - \sum_{l\neq i} \bar{b}_l \mu_l \left[\tilde{\phi}_{lt} - \left(1 - 2^{-2\kappa_j^{(l)}}\right)\left(\tilde{\phi}_{lt} + \psi^{(l)}_{jt}\right)\right] \\
& - (\bar{b}_i \mu_i + \gamma)\left[\tilde{\phi}_{it} - \left(1 - 2^{-2\kappa_j^{(i)}}\right)\left(\tilde{\phi}_{it} + \psi^{(i)}_{jt}\right)\right] \\
=& \bar{a}\left[2^{-2\kappa_j^{(q)}}\tilde{q}_t - \left(1 - 2^{-2\kappa_j^{(q)}}\right)\psi^{(q)}_{jt}\right] \\
& - \sum_{l\neq i} \bar{b}_l \mu_l \left[2^{-2\kappa_j^{(l)}}\tilde{\phi}_{lt} - \left(1 - 2^{-2\kappa_j^{(l)}}\right)\psi^{(l)}_{jt}\right] \\
& - (\bar{b}_i \mu_i + \gamma)\left[2^{-2\kappa_j^{(i)}}\tilde{\phi}_{it} - \left(1 - 2^{-2\kappa_j^{(i)}}\right)\psi^{(i)}_{jt}\right]
\end{aligned}$$

then noting that independence implies that all cross terms have expected value zero, the mean squared error can be expressed

$$\begin{aligned}
E_{jt}\left[\left(\tilde{p}^{\diamond}_{jt} - \tilde{p}^{*}_{jt}\right)^2\right] =& \bar{a}^2\left[2^{-4\kappa_j^{(q)}}\sigma_q^2 + \left(1 - 2^{-2\kappa_j^{(q)}}\right)^2 \sigma^2_{\psi_j^{(q)}}\right] \\
& + \sum_{l=1}^{I} \bar{b}_l^2 \mu_l^2 \left[2^{-4\kappa_j^{(l)}}\sigma^2_{\phi_l} + \left(1 - 2^{-2\kappa_j^{(l)}}\right)^2 \sigma^2_{\psi_j^{(l)}}\right] \\
& + (\bar{b}_i \mu_i + \gamma)^2 \left[2^{-4\kappa_j^{(i)}}\sigma^2_{\phi_i} + \left(1 - 2^{-2\kappa_j^{(i)}}\right)^2 \sigma^2_{\psi_j^{(i)}}\right]
\end{aligned}$$

then using result (4) from *D3: Expressions for Mutual Information*, it can be finally written

$$E_{jt}\left[\left(\tilde{p}^{\diamond}_{jt} - \tilde{p}^{*}_{jt}\right)^2\right] = \bar{a}^2 2^{-2\kappa_j^{(q)}}\sigma_q^2 + \sum_{l=1}^{I} \bar{b}_l^2 \mu_l^2 2^{-2\kappa_j^{(l)}}\sigma^2_{\phi_l} + (\bar{b}_i \mu_i + \gamma)^2 2^{-2\kappa_j^{(i)}}\sigma^2_{\phi_i}$$

E-2.3: The Attention Problem

The firm's attention problem can now be fully specified

$$\min_{\{\kappa_j^{(\omega)}\}_{\omega\in\Omega}} \sum_{\omega\in\Omega} \left(\bar{\kappa}_j^{(\omega)}\right)^2 2^{-2\kappa_j^{(\omega)}} \qquad ; \qquad \bar{\kappa}_j^{(\omega)} = \begin{cases} \bar{a}\sigma_q & \omega = q \\ \bar{b}_l\mu_l\sigma_{\phi_l} & \omega = l \neq i \\ (\bar{b}_i\mu_i + \gamma)\sigma_{\phi_i} & \omega = i \end{cases}$$

such that $\sum_{\omega\in\Omega} \kappa_j^{(\omega)} \leq \kappa$ and $\kappa_j^{(\omega)} \geq 0$. This is a constrained optimization problem and can be represented as a Lagrangian (where the constraint is assumed to be binding, since in any optimum firms will use all available attention)

$$\mathcal{L} = \sum_{\omega\in\Omega} \left(\bar{\kappa}_j^{(\omega)}\right)^2 2^{-2\kappa_j^{(\omega)}} - \lambda \left[\sum_{\omega\in\Omega} \kappa_j^{(\omega)} - \kappa\right]$$

Assuming an interior solution, the $|\Omega|$ first-order conditions for an optimum are

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\kappa_j^{(\omega)}} = 0 &= \left(\bar{\kappa}_j^{(\omega)}\right)^2 2^{-2\kappa_j^{(\omega)}}(-2\ln 2) - \lambda \\ 2^{2\kappa_j^{(\omega)}} &= \frac{\left(\bar{\kappa}_j^{(\omega)}\right)^2(-2\ln 2)}{\lambda} \\ \kappa_j^{(\omega)} &= \frac{1}{2}\log_2\left(\frac{-2\ln 2}{\lambda}\right) + \frac{1}{2}\log_2\left[\left(\bar{\kappa}_j^{(\omega)}\right)^2\right] \end{aligned}$$

Define $\hat{\kappa}_k^{(\omega)} = \frac{1}{2}\log_2\left[\left(\bar{\kappa}_j^{(\omega)}\right)^2\right]$ to ease notation and use the first condition to substitute out the Lagrange multiplier

$$\frac{1}{2}\log_2\left(\frac{-2\ln 2}{\lambda}\right) = \kappa_j^{(\omega_1)} - \hat{\kappa}_j^{(\omega_1)}$$

Then use this in the remaining $|\Omega| - 1$ conditions to get:

$$\begin{aligned} \kappa_j^{(\omega_k)} &= \kappa_j^{(\omega_1)} - \hat{\kappa}_j^{(\omega_1)} + \hat{\kappa}_j^{(\omega_k)} \\ -\kappa_j^{(\omega_1)} + \kappa_j^{(\omega_k)} &= -\hat{\kappa}_j^{(\omega_1)} + \hat{\kappa}_j^{(\omega_k)} \qquad\qquad k = 2, \ldots, |\Omega| \end{aligned}$$

Including the constraint $\sum_{\omega \in \Omega} \kappa_j^{(\omega)} = \kappa$ there are $|\Omega|$ equations and $|\Omega|$ unknowns. This can be written in the following linear system:

$$\begin{bmatrix} -1 & 1 & 0 & \cdots & & 0 \\ -1 & 0 & 1 & \cdots & & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ -1 & \cdots & & 0 & 1 & 0 \\ -1 & \cdots & & & 0 & 1 \\ 1 & 1 & \cdots & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \kappa_j^{(\omega_1)} \\ \kappa_j^{(\omega_2)} \\ \vdots \\ \kappa_j^{(\omega_{|\Omega|-1})} \\ \kappa_j^{(\omega_{|\Omega|})} \end{bmatrix} = \begin{bmatrix} -\hat{\kappa}_j^{(\omega_1)} + \hat{\kappa}_j^{(\omega_2)} \\ -\hat{\kappa}_j^{(\omega_1)} + \hat{\kappa}_j^{(\omega_3)} \\ \vdots \\ -\hat{\kappa}_j^{(\omega_1)} + \hat{\kappa}_j^{(\omega_{|\Omega|})} \\ \kappa \end{bmatrix}$$

This can be solved using the following steps:

1. Multiply row 1 by -1
2. For rows 2 through $|\Omega|$, iteratively add the previous row and multiply by -1
3. For rows $l = 1, \ldots, |\Omega| - 1$, substract l times the l^{th} for from row $|\Omega|$.
4. Divide row $|\Omega|$ by $|\Omega|$
5. For rows $l = |\Omega| - 1, \ldots, 1$, add the $l + 1^{\text{th}}$ row
6. Simplify

This process yields optimal interior allocations that can be expressed as

$$\kappa_j^{(\omega)*} = |\Omega|^{-1} \left[\kappa - \sum_{\omega' \neq \omega} \hat{\kappa}_j^{(\omega')} + (|\Omega| - 1) \hat{\kappa}_j^{(\omega)} \right] \qquad\qquad \omega \in \Omega$$

We can abuse notation to take into account the corner conditions

$$\kappa_j^{(\omega)^*} = \begin{cases} \kappa & \kappa_j^{(\omega)^*} > \kappa \\ \kappa_j^{(\omega)^*} & \kappa_j^{(\omega)^*} \in [0, \kappa] \\ 0 & \kappa_j^{(\omega)^*} < 0 \end{cases}$$

The allocations can be rewritten

$$\begin{aligned} \kappa_j^{(\omega)^*} &= |\Omega|^{-1} \left[\kappa - \sum_{\omega' \in \Omega} \hat{\kappa}_j^{(\omega')} + |\Omega| \hat{\kappa}_j^{(\omega)} \right] \\ &= |\Omega|^{-1} \left[\frac{1}{2} \log_2 \left(2^{2\kappa} \right) - \sum_{\omega' \in \Omega} \frac{1}{2} \log_2 \left[\left(\bar{\kappa}_j^{(\omega')} \right)^2 \right] + |\Omega| \frac{1}{2} \log_2 \left[\left(\kappa_j^{(\omega)} \right)^2 \right] \right] \end{aligned}$$

defining $\bar{\kappa} = \frac{\kappa}{|\Omega|}$ and $\bar{\kappa}_j = \left[\prod_{\omega' \in \Omega} \bar{\kappa}_j^{(\omega')} \right]^{\frac{1}{|\Omega|}}$, we have the final expression for the optimal allocation of attention

$$\kappa_j^{(\omega)^*} = \log_2 2^{\bar{\kappa}} + \log_2 \bar{\kappa}_j^{(\omega)} - \log_2 \bar{\kappa}_j \qquad\qquad \omega \in \Omega$$

Note that it is straightforward that

$$1 - 2^{-2\kappa_j^{(\omega)^*}} = 1 - \left(2^{-2\bar{\kappa}} \right) \left(\bar{\kappa}_j^{(\omega)} \right)^{-2} \bar{\kappa}_j^2$$

E-2.4: Verifying the Guess

Returning to the optimal imperfect-information pricing rule

$$\begin{aligned}
\tilde{p}_{jt}^{*} =& \bar{a}\left(1-2^{-2\kappa_j^{(q)^*}}\right)\left(\tilde{q}_t-\psi_{jt}^{(q)}\right) \\
&-\sum_{l\neq i}\bar{b}_l\mu_l\left(1-2^{-2\kappa_j^{(l)^*}}\right)\left(\tilde{\phi}_{lt}-\psi_{jt}^{(l)}\right) \\
&-\left(\bar{b}_i\mu_i+\gamma\right)\left(1-2^{-2\kappa_j^{(i)^*}}\right)\left(\tilde{\phi}_{it}-\psi_{jt}^{(i)}\right)
\end{aligned}$$

Integrating over all firms, applying symmetry to within-industry firms' optimal attention allocations, and applying the log-linear approximation to the aggregate price yields

$$\begin{aligned}
\tilde{p}_{t}^{*} =& \int_J \bar{a}\left(1-2^{-2\kappa_i^{(q)^*}}\right)\tilde{q}_t dj \\
&-\int_J\left[\sum_{l\neq i}\bar{b}_l\mu_l\left(1-2^{-2\kappa_i^{(l)^*}}\right)\tilde{\phi}_{lt}\right]dj \\
&-\int_J\left(\bar{b}_i\mu_i+\gamma\right)\left(1-2^{-2\kappa_i^{(i)^*}}\right)\tilde{\phi}_{it}dj
\end{aligned}$$

notice that the noise variables are mean zero and firm-specific so that the integral with respect to a continuum of firms is equal to zero. Uhlig (1996)

Then we have

$$
\begin{aligned}
\tilde{p}_t^* &= \left[\sum_{i=1}^{I} \mu_i \bar{a}\left(1-2^{-2\kappa_i^{(q)*}}\right)\right]\tilde{q}_t - \sum_{i=1}^{I}\mu_i\left[\sum_{l\neq i}\bar{b}_l\mu_l\left(1-2^{-2\kappa_i^{(l)*}}\right)\tilde{\phi}_{lt}\right] \\
&\quad - \sum_{i=1}^{I}\mu_i(\bar{b}_i\mu_i+\gamma)\left(1-2^{-2\kappa_i^{(i)*}}\right)\tilde{\phi}_{it} \\
&= \left[\sum_{i=1}^{I} \mu_i \bar{a}\left(1-2^{-2\kappa_i^{(q)*}}\right)\right]\tilde{q}_t - \frac{\gamma}{\zeta}\sum_{i=1}^{I}\mu_i\left[\sum_{l=1}^{I}(1-\zeta)b_l\mu_l\left(1-2^{-2\kappa_i^{(l)*}}\right)\tilde{\phi}_{lt}\right] \\
&\quad - \frac{\gamma}{\zeta}\sum_{i=1}^{I}\zeta\mu_i\left(1-2^{-2\kappa_i^{(i)*}}\right)\tilde{\phi}_{it} \\
&= \left[\sum_{i=1}^{I}\mu_i\left[(1-\zeta)a+\zeta\right]\left(1-2^{-2\kappa_i^{(q)*}}\right)\right]\tilde{q}_t - \frac{\gamma}{\zeta}\sum_{l=1}^{I}\left[(1-\zeta)b_l\sum_{i=1}^{I}\mu_i\left(1-2^{-2\kappa_i^{(l)*}}\right)\right]\mu_l\tilde{\phi}_{lt} \\
&\quad - \frac{\gamma}{\zeta}\sum_{l=1}^{I}\zeta\left(1-2^{-2\kappa_l^{(l)*}}\right)\mu_l\tilde{\phi}_{lt} \\
&= \left[\sum_{i=1}^{I}\mu_i\left[(1-\zeta)a+\zeta\right]\left(1-2^{-2\kappa_i^{(q)*}}\right)\right]\tilde{q}_t - \frac{\gamma}{\zeta}\sum_{l=1}^{I}\left[\sum_{i=1}^{I}w_{li}\left(1-2^{-2\kappa_i^{(l)*}}\right)\right]\mu_l\tilde{\phi}_{lt}
\end{aligned}
$$

where $w_{il} = (1-\zeta)b_l\mu_i + \zeta\mathbf{1}(l=i)$ and $\mathbf{1}(l=i)$ is the indicator function that takes the value 1 if $l=i$ and is 0 otherwise. This verifies the guess with

$$
\begin{aligned}
a &= \left[\sum_{i=1}^{I}\mu_i\left[(1-\zeta)a+\zeta\right]\left(1-2^{-2\kappa_i^{(q)*}}\right)\right] \\
b_l &= \left[\sum_{i=1}^{I}\left[(1-\zeta)b_l\mu_i+\zeta\mathbf{1}(l=i)\right]\left(1-2^{-2\kappa_i^{(l)*}}\right)\right]
\end{aligned}
$$

these can be rewritten to emphasize the effect of the parameter of strategic complementarities

$$
\begin{aligned}
a &= (1-\zeta)a\sum_{i=1}^{I}\mu_i\left(1-2^{-2\kappa_i^{(q)*}}\right) + \zeta\sum_{l=1}^{I}\mu_l\left(1-2^{-2\kappa_l^{(q)*}}\right) \\
b_l &= (1-\zeta)b_l\sum_{i=1}^{I}\mu_i\left(1-2^{-2\kappa_i^{(l)*}}\right) + \zeta\left(1-2^{-2\kappa_l^{(l)*}}\right)
\end{aligned}
$$

Appendix F: Notation

8.14 F-1: Parameters

Indices

- $h \in H$ index for households
- μ_H measure for households with $\mu_H(H) = 1$
- $j \in J$ index for firms
- μ_J measure for households with $\mu_J(J) = 1$
- $\mu_i \equiv \mu_J(J_i)$ convenience notation
- $i \in \{1, \ldots, I\}$ index for sectors
- $\{J_1, \ldots, J_I\}$ partition on firms induced by sectors

Households

- σ: coefficient of relative risk aversion; inverse of the elasticity of intertemporal substitution parameter
- ε: inverse of Frisch elasticity of labor supply "measures the substitution effect of a change in the wage rate on labor supply." Comes from the derived Household optimization equation $w_t = c_t^\sigma n_t^\varepsilon$, so that $n_t = w_t^{\frac{1}{\varepsilon}} c_t^{\frac{\sigma}{\varepsilon}}$.
- $r \in [0, 1)$: within-industry generalized mean exponent
- $p \in [0, 1)$: between-industry generalized mean exponent
- $\eta = \frac{1}{1-r} \in [1, \infty)$: within-industry elasticity of substitution; measure of market power

- $\rho = \frac{1}{1-p} \in [1, \infty)$: between-industry elasticity of substitution; measure of trade linkages

Equilibrium

- $\alpha = \frac{1}{1+\rho\varepsilon}$ parameterizes strategic complementarities specifically arising from heterogeneous information, see Angeletos and La'O (2010).
- $\gamma = \alpha(1+\varepsilon)$
- $\zeta = \alpha(\sigma+\varepsilon)$ is the typical New Keynesian parameter governing strategic complementarities generally (see Morris and Shin, 2002, Woodford (2003) chapter 3, and Mankiw and Reis (2010)) and also related to the degree of "real rigidities" (see Ball and Romer, 1990).

8.15 F-2: Stochastic Processes

- $\Omega = \{\{q_t\}, \{\phi_{1t}\}, \cdots, \{\phi_{It}\}\}$ is an ordered tuple gathering all stochastic processes and indexed by ω.
- ω_l is the l^{th} item in Ω; for example $\omega_1 \equiv q$.

Fundamentals

- $q_t \overset{iid}{\sim} N(0, \sigma_q^2)$: nominal aggregate demand
- $\phi_{it} \overset{iid}{\sim} N(0, \sigma_{\phi_i}^2)$ idiosyncratic productivity shocks

Signals

- $s_{jt}^{(q)} = \tilde{q}_t + \psi_{jt}^{(q)}$ is the signal to firm j related to aggregate conditions. $s_{jt}^{(q)} \sim N(\tilde{q}_t, \sigma_q^2 + \sigma^2_{\psi_j^{(q)}})$
- $s_{jt}^{(l)} = \tilde{\phi}_{;t} + \psi_{jt}^{(l)}$ is the signal to firm j related to the productivity shock to industry l. $s_{jt}^{(l)} \sim N(\tilde{\phi}_{lt}, \sigma^2_{\phi_l} + \sigma^2_{\psi_j^{(l)}})$

8.16 F-3: Information

- κ represents the Shannon capacity of a channel, measured in bits. This term is also used for the specific parameter describing total capacity available to agents.
- $\kappa_j^{(\omega)} \equiv \mathcal{I}(\omega, s_j^{(\omega)})$ represents the information capacity allocated by firm j to stochastic process ω
- $\kappa_j^{(\omega)^*}$ represents the optimal allocated capacity by firm j
- $\kappa_i^{(\omega)^*}$ represents the optimal capacity by any firm in sector i allocated to stochastic process ω; requires appealing to symmetry

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The Network Origins of Aggregate Fluctuations. *Econometrica 80*(5), 1977–2016. 00227.

Angeletos, G.-M. and J. La'O (2010). Noisy Business Cycles. *NBER Macroeconomics Annual 2009 24*, 319–378. 00081.

Ball, L. and D. Romer (1990, April). Real Rigidities and the Non-Neutrality of Money. *The Review of Economic Studies 57*(2), 183–203. 00723.

Basu, S. (1995, June). Intermediate Goods and Business Cycles: Implications for Productivity and Welfare. *The American Economic Review 85*(3), 512–531. 00330.

Bils, M. and P. Klenow (2004, October). Some Evidence on the Importance of Sticky Prices. *Journal of Political Economy 112*(5), 947–985. 00000 ArticleType: research-article / Full publication date: October 2004 / Copyright © 2004 The University of Chicago Press.

Boivin, J., M. P. Giannoni, and I. Mihov (2009, March). Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data. *The American Economic Review 99*(1), 350–384. 00000.

Bouakez, H., E. Cardia, and F. J. Ruge-Murcia (2009, September). Sectoral Price Rigidity and Aggregate Dynamics. SSRN Scholarly Paper ID 1548020, Social Science Research Network, Rochester, NY. 00010.

Calvo, G. A. (1983, September). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics 12*(3), 383–398.

Carvalho, C. V. d. and J. W. Lee (2011, November). Sectoral Price Facts in a Sticky-Price Model. Departmental Working Paper, Rutgers University, Department of Economics.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2000, September). Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem? *Econometrica 68*(5), 1151–1179. ArticleType: research-article / Full publication date: Sep., 2000 / Copyright © 2000 The Econometric Society.

Cover, T. M. and J. A. Thomas (2006, July). *Elements of Information Theory*. John Wiley & Sons. 00076.

De Graeve, F. and K. Walentin (2014, September). Refining Stylized Facts from Factor Models of Inflation. *Journal of Applied Econometrics*, n/a–n/a. 00000.

Dixit, A. K. and J. E. Stiglitz (1977, June). Monopolistic Competition and Optimum Product Diversity. *The American Economic Review 67*(3), 297–308. 08233.

Dotsey, M., R. G. King, and A. L. Wolman (1999, May). State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output. *The Quarterly Journal of Economics 114*(2), 655–690. 00628.

Friedman, M. (1968, March). The Role of Monetary Policy. *The American Economic Review 58*(1), 1–17. 00006.

Golosov, M. and R. E. Lucas Jr. (2007, April). Menu Costs and Phillips Curves. *Journal of Political Economy 115*(2), 171–199. 00002.

Klenow, P. J. and O. Kryvtsov (2008, August). State-Dependent or Time-Dependent Pricing: Does it Matter for Recent U.S. Inflation? *The Quarterly Journal of Economics 123*(3), 863–904. 00543.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory 4*(2), 103–124. 04145.

Mankiw, N. G. and R. Reis (2002, November). Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *The Quarterly Journal of Economics 117*(4), 1295–1328. 01501 ArticleType: research-article / Full publication date: Nov., 2002 / Copyright © 2002 Oxford University Press.

Mankiw, N. G. and R. Reis (2010). Imperfect Information and Aggregate Supply. Handbook of Monetary Economics, Elsevier. 00057.

Maćkowiak, B., E. Moench, and M. Wiederholt (2009, October). Sectoral price data and models of price setting. *Journal of Monetary Economics 56, Supplement*, S78–S99. 00045.

Maćkowiak, B. and M. Wiederholt (2009, June). Optimal Sticky Prices under Rational Inattention. *The American Economic Review 99*(3), 769–803. 00324.

Morris, S. and H. S. Shin (2002, December). Social Value of Public Information. *The American Economic Review 92*(5), 1521–1534. 01056.

Phelps, E. S. (1968). Money-Wage Dynamics and Labor-Market Equilibrium. *Journal of Political Economy 76*. 00000.

Shannon, C. E. (1948, July). A Mathematical Theory of Communication. *Bell System Technical Journal 27*(3), 379–423. 00004.

Sims, C. (2010). Rational Inattention and Monetary Economics. Handbook of Monetary Economics, Elsevier. 00065.

Sims, C. A. (1998, December). Stickiness. *Carnegie-Rochester Conference Series on Public Policy 49*, 317–356. 00204.

Sims, C. A. (2003, April). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690. 01010.

Sims, C. A. (2005). Rational inattention: a research agenda. Discussion Paper Series 1: Economic Studies 2005,34, Deutsche Bundesbank, Research Centre.

Townsend, R. M. (1983, August). Forecasting the Forecasts of Others. *Journal of Political Economy 91*(4), 546–588. 00388.

Uhlig, H. (1996, June). A Law of Large Numbers for Large Economies. *Economic Theory 8*(1), 41–50.

Woodford, M. (2001, December). Imperfect Common Knowledge and the Effects of Monetary Policy. Working Paper 8673, National Bureau of Economic Research. 00453.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press. 00022.