Chapter 6

Informational Inertia and Price-Setting

Much of the work on imperfect information and information choice in macroeconomics has been done in the context of price-setting models. One of the most fundamental questions macroeconomics faces is why monetary policy appears to have real effects on the economy. In frictionless models, prices are typically proportional to the money supply, so that money simply determines the units of account, but not real output. The question put differently is: Why is the covariance of prices with the money supply so much lower than what standard models predict? Since information choice governs the covariance between choice variables (prices) and states (the money supply), it is a tool well-suited to address this problem.

The chapter proceeds by describing a sequence of models. Each builds on the ideas of the one before. A central question in monetary economics is how the supply of money affects real output. Lucas (1972) sets the stage for this literature by arguing that the key is imperfect information about aggregate variables such as the price level and the money supply. A problem with the Lucas model is that it does not explain why output has a *delayed* reaction to changes in the money supply. Woodford (2002) argues that a combination of heterogeneous information and coordination motives in price setting can create inertia in prices. Since money supply and price levels are published frequently, imperfect, heterogeneous information presumably arises because accessing or processing this public information is costly. Reis (2006) investigates the predictions of a model with fixed costs of updating beliefs. But firms that do not update their beliefs frequently should not react promptly to any shocks.

Yet, they appear to react aggressively to firm-specific shocks, just not to monetary shocks. In Maćkowiak and Wiederholt (2009b), firms optimally choose to learn more about firm-specific than monetary shocks because the firm-specific shocks generate more volatility in their optimal price. Finally, Woodford (2008) incorporates both the fixed-cost friction in Reis (2006) and the rational inattention friction of Maćkowiak and Wiederholt (2009b) into one framework in order to assess their relative importance.

While the specific message of each of these papers about the dynamics of price setting and efficacy of monetary policy are mostly of interest to people who work in this area, they could describe any setting where there is a dynamic problem with strategic interactions in actions and information choice. As such, the chapter explores two broad questions that extend beyond the application to monetary policy. *Question #1: Why do some shocks seem to matter, while others do not?* Prices of goods, or equity, production and investment decisions, all respond differently to different types of shocks. Rational inattention offers one potential explanation for why that might be. *Question #2: Why do actions exhibit inertia?* Investors are slow to rebalance their portfolios. Consumers are slow to adjust their consumption in response to income shocks. Some industries lag the business cycle more than others. One explanation could be that the incentives to process aggregate information may differ according to the type of information, or according the to agent and the nature of the shocks they face.

6.1 Lucas-Phelps Model

The original Lucas (1972) model has overlapping generations, a more general utility, two islands and uncertainty about both the supply of money that old agents have and the number of young agents working on each island. This model is difficult to work with in its original form. We will examine instead a simplified version of that model. It misses some of the micro-foundations, but will give a clear picture of the kind of informational problem agents face.

There is a continuum of goods in the economy, each produced by a single representative producer i. Each producer can transform labor L_i, one-for-one into good-i output Y_i

$$Y_i = L_i. \tag{6.1}$$

Utility is defined over the consumption of a composite good C and labor

$$U_i = C_i - \frac{1}{\gamma} L_i^{\gamma} \tag{6.2}$$

where $\gamma > 1$. The composite good C_i is constructed so as to deliver an aggregate demand for good i that depends on aggregate income Y, good i's price P_i, relative to the price of the consumption aggregate, and a random, mean-zero shock to the preference for good i, z_i:

$$Y_i^d = Y \left(\frac{P_i}{P} \right)^{-\eta} \exp(z_i), \tag{6.3}$$

where aggregate income is defined such that $\ln(Y) \equiv \int_i \ln(Y_i) di$. Aggregate demand for goods depends on the real money supply M and the aggregate price level P

$$Y = M/P \tag{6.4}$$

where the price level is defined such that$\ln(P) = \int_i \ln(P_i) di$. Specifying demand directly like this is obviously a shortcut. (See Lorenzoni (2009) for a micro-founded model that delivers this form of aggregate demand.) Finally, an individual's budget constraint is

$$PC_i = P_i Y_i. \tag{6.5}$$

Equilibrium An equilibrium is a set of utility-maximizing labor L_i and consumption C_i choices that respect the budget constraint and prices that equate demand and supply. Substituting the budget constraint (6.5) and the production function (6.1) into the individual's utility (6.2) delivers $E[U_i] = E[P_i L_i / P] - (1/\gamma) L_i^{\gamma}$.The first order condition in L_i reveals that the optimal choice of labor is

$$L_i = E \left[\frac{P_i}{P} \right]^{1/(\gamma - 1)}. \tag{6.6}$$

Equilibrium Prices with Full Information With full information, we can drop the expectation operator in (6.6). To simplify the analysis, transform the problem into logarithms. Let lower case variables denote logs. For example $l_i \equiv \ln(L_i)$.

The log of supply (equation 6.6) is $1/(\gamma - 1)(p_i - p)$. The log of demand (equation 6.3) is $y - \eta(p_i - p) + z_i$. Equating these two expressions and rearranging delivers the equilibrium price:

$$p_i = p + \frac{\gamma - 1}{1 + \eta(\gamma - 1)}(y + z_i). \tag{6.7}$$

Since the log aggregate price is defined to be an average of the log individual prices,

$$p = \int p_i di = p + \frac{\gamma - 1}{1 + \eta(\gamma - 1)}(y + 0). \tag{6.8}$$

Subtracting p on both sides leaves log output $y = 0$, meaning that aggregate output Y is always $e^0 = 1$. Using the aggregate demand equation (6.4), it also tells us that $M = P$. In short, money is neutral. A large supply of money only affects prices, not output.

Equilibrium Prices with Incomplete Information The key assumption that causes money to have real effects is that agents do not know the aggregate price level. Suppose that, when choosing how much to produce or work, individuals can see the price of their good p_i, but the aggregate price level p, the money supply m and the good-specific demand shock z_i are all unobserved. They believe these shocks to be normally distributed: $m \sim (E[m], V_m)$ and $z_i \sim N(0, V_z)$. If p is a linear function of m and z_i, then p will have an unconditional (not conditioning on observing p_i) normal distribution: $p \sim N(E[p], V_p)$, for some values $E[p]$ and V_p. After production takes place, the aggregate price level and money supply are revealed, so that consumers' demand can depend on p and m.

Along with the limited-information assumption, Lucas makes a simplifying assumption – certainty equivalence. He assumes that, when choosing their optimal labor supply, agents form an expectation $E_i[p_i - p] \equiv E[p_i - p|p_i]$ and then treat that expectation as if it were the truth. Since the log optimal labor supply (log of equation 6.6) with full information is $1/(\gamma - 1)(p_i - p)$, the certainty equivalent of optimal labor is

$$l_i = \frac{1}{\gamma - 1} E_i[p_i - p] \tag{6.9}$$

The interpretation is that agents work harder when they believe that the relative price of their good $(p_i - p)$ is high. Money has real effects because when money is abundant, most agents observe a high price for their good. But each agent cannot tell whether their price is high because their good's relative price is high or because the aggregate price level is high. So, they place some probability on each cause. If their expected relative price rises, they work harder, which increases output.

The equilibrium price is determined by equating supply (6.9) with demand. Demand is the same as in the full-information model (the log of equation 6.3). Using (6.4) to substitute out y,

$$\frac{1}{\gamma - 1}(p_i - E_i[p]) = m - p - \eta(p_i - p) + z_i. \tag{6.10}$$

Note that this is a linear relationship between p and p_i, with normally distributed noise $m + z_i$. Since Bayes' law for normal variables is also a linear relationship, $E_i[p]$ will be linear as well. Thus, p_i must be equal to a linear function of p, plus noise. In other words, there are some coefficients a and b such that $p_i = a + bp + \epsilon_i$ where $\epsilon_i \sim N(0, V_e)$. Since we know that p is always the average log price, $p = \int p_i di = a + bp + 0$. This can only hold for every aggregate price if $a = 0$ and $b = 1$. Thus,

$$p_i = p + \epsilon_i \qquad \epsilon_i \sim N(0, V_e).$$

This is useful because it tells us that the price of an individual good is an unbiased signal about the aggregate price level.

Next, use Bayes' law to combine this information in the individual good price with the aggregate price to get a conditional expectation $E_i[p]$. As explained in chapter 2, the conditional expectation (the posterior belief) is a weighted sum of the prior and signal, where the weights are the relative precisions of each. Thus,

$$E_i[p] = \frac{V_e}{V_p + V_e} E[p] + \frac{V_p}{V_p + V_e} p_i. \tag{6.11}$$

The next step is to substitute for $E_i[p]$ in the labor supply equation (6.9). Define $\alpha \equiv V_e/(V_p + V_e) \cdot 1/(\gamma - 1)$ and collect terms to get

$$l_i = \alpha\left(p_i - E[p]\right). \tag{6.12}$$

Averaging over i delivers aggregate production ($\int_i l_i di = y$) and using the aggregate demand relationship $y = m - p$ yields

$$m - p = \alpha\left(p - E[p]\right). \tag{6.13}$$

Taking the unconditional expectation of both sides of (6.13) tells us that $E[m] - E[p] = 0$. Thus, $E[p] = E[m]$. The unconditional expected price level is the expected money supply. This parallels the result in the full-information model. When m and p were known, we had

$m = p$.

Substituting in $E[m]$ for $E[p]$ in (6.13) and rearranging, we can express the aggregate price level as a function of the expected and actual money supply

$$p = \frac{\alpha}{1+\alpha}E[m] + \frac{1}{1+\alpha}m. \tag{6.14}$$

This captures the idea that inflation can be driven by changes in monetary policy m or changes in expectations $E[m]$.

Output depends on the difference between actual and expected money supply. Using the aggregate demand relationship $y = m - p$ again reveals that

$$y = \frac{\alpha}{1+\alpha}(m - E[m]). \tag{6.15}$$

The main result of the Lucas-Phelps model is that money supply surprises cause output to rise. This is important because it creates a role for monetary policy to have real effects on economic activity.

6.2 A Recipe for Inertia

The Lucas island model's results are important because they describe how monetary policy can be used to manage the real economy. They embody an expectations-augmented Phillips' curve – the idea that when inflation (or money growth) is higher than expected, the economy's output will be above trend. The problem with Lucas' explanation is that once producers produce and turn into consumers, they observe all posted prices. If the aggregate price level from period t is known in period $t+1$, then the real effects of monetary policy should be only one-period-lived. Yet, in the data, monetary policy has its greatest effect on output 6 quarters after the monetary shock, and after 10 quarters, this effect is still 1/3rd the size of the peak effect (Woodford, 2002).

Of course, one could argue about the length of periods in the Lucas model. But with all the statistics released daily through the prices in financial markets and published quarterly by government bureaus, it is difficult to defend the idea that aggregate price information is not available more than two years after the fact.

Woodford (2002) argues that public information about the aggregate price level is readily available. The friction is that firms cannot process all the information they observe. They can observe it freely, but figuring out what it means and how to use it to set their good's price requires capacity, which is in limited supply. This limited capacity is modeled as an exogenous noisy signal that each firm gets about the state of the economy, whose signal noise is independent of other firms' signals. That noisy signal informs firms about the current and past periods' states so that beliefs converge slowly to the truth.

Woodford makes a second important change to the Lucas model: He introduces a complementarity into the payoff structure. It gives firms an incentive to coordinate price-setting. In order to coordinate effectively, they want to forecast other firms' beliefs. Learning about what others know is a much slower process than learning the value of an exogenous state variable. This is because higher-order beliefs are much more uncertain that first-order beliefs (see section 4.1). This slow updating helps monetary policy shocks have long-lived effects. We investigate a simplified version of that model where log GDP is a random walk.

There is a continuum of firms, indexed by i. Each firm optimally chooses its price p_{it} to maximize its expected profit. Woodford (2002) shows that a standard profit maximization objective can be approximated by the following second-order utility function

$$U = E_{it}\left[\bar{u} - r(p_{it} - p_t)^2 - (1-r)(p_{it} - q_t)^2\right] \tag{6.16}$$

where $p_t \equiv \int p_{it} di$ and q_t are the (log) time-t aggregate price level and nominal GDP and E_{it} denotes firm i's expectations, based on its time-t information set. The parameter r measures the degree of complementarity in price-setting. The higher r is, the more aggregate prices matter relative to fundamental shocks, and the stronger the complementarity in price-setting is.

Log nominal GDP is a random walk with a known drift g and an unknown innovation u_t:

$$q_t = g + q_{t-1} + u_t \qquad u_t \sim N(0, \sigma_u^2). \tag{6.17}$$

Finally, the information each firm receives each period is a noisy signal of current nominal GDP

$$z_{it} = q_t + v_{it} \tag{6.18}$$

where v_{it} is i.i.d.$N(0, \sigma_v^2)$ across individuals and over time. More capacity to process information corresponds to a lower σ_v.

Characterizing the model's solution The first-order condition for an optimum reveals that $p_{it} = E_{it}[p_t] + (1-r)E_{it}[q_t]$. If we average over all firms, we get the aggregate price level as a function of the average beliefs $\bar{E}_t[p_t]$ and $\bar{E}_t[q_t]$,

$$p_t = r\bar{E}_t[p_t] + (1-r)\bar{E}_t[q_t]. \tag{6.19}$$

This model describes a form of a coordination game with heterogeneous expectations. Just like in chapter 4.1, the aggregate action (p_t in this case) can be described as an infinite sum of higher-order expectations. To see this, recursively substitute for p_t on the right side of the equality to get $p_t = \sum_{k=1}^{\infty}(1-r)r^{(k-1)}q_t^{(k)}$, where the superscript (k) represents the k^{th}-order average expectation. For example, $q_t^{(1)} = \bar{E}_t[q_t]$ is the average belief about q_t, while $q_t^{(2)} = \bar{E}_t[q_t^{(1)}]$ is the average belief about the average belief of q_t, and so forth. Working with this infinite sum is intractable. Morris and Shin (2002) avoid this problem by conjecturing and then verifying a symmetric strategy for firms. Woodford (2002) instead conjectures that q_t and p_t are sufficient state variables and that they have a linear law of motion.

Define $x_t \equiv [q_t\ p_t]'$ to be the 2×1 vector of state variables. Then, the conjectured law of motion is

$$x_{t+1} = c + Mx_t + mu_{t+1}, \tag{6.20}$$

where the unknown coefficients are c (2×1), M (2×2), and m (2×1) and where u_{t+1} is the innovation to GDP in (6.17).

Note that (6.18) and (6.20) comprise the observation and state equations of a Kalman filter. Using the equation (2.8) from chapter 2, we can express agent i's expected value of state x at time t as

$$E_{it}[x_t] = c + ME_{i(t-1)}[x_{t-1}] + k_t(z_{it} - E_{it}[q_t]), \tag{6.21}$$

where k_t is the Kalman gain. Integrating this expectation over agents and using the fact that the signals z_{it} have mean q_t gives the average expectation

$$\bar{E}_t[x_t] = c + M\bar{E}_{t-1}[x_{t-1}] + k_t(q_t - \bar{E}_t[q_t]). \tag{6.22}$$

Note that the Kalman gain is identical across agents because we have assumed that all have the same prior and signal variances.

The next step is to solve for the unknown coefficients. The assumption that GDP is a random walk (equation 6.17) tells us what the coefficients in the first row of (6.22) are: $c(1) = g$, $M(1,1) = 1$ and $M(1,2) = 0$. Equation (6.22) has two rows, which we can write out separately. The first is $\bar{E}_t[q_t] = g + \bar{E}_{t-1}[q_{t-1}] + k_t(1)(q_t - \bar{E}_t[q_t])$. The second row is $\bar{E}_t[p_t] = c(2) + M(2,1)\bar{E}_{t-1}[q_{t-1}] + M(2,2)\bar{E}_{t-1}[p_{t-1}]) + k_t(2)(q_t - \bar{E}_t[q_t])$. To determine the unknown coefficients, substitute these two expressions in for $\bar{E}_t[q_t]$ and $\bar{E}_t[p_t]$ in (6.19) to get

$$\begin{aligned} p_t &= r\left[c_2 + M(2,1)\bar{E}_{t-1}[q_{t-1}] + M(2,2)\bar{E}_{t-1}[p_{t-1}]\right] + (1-r)\left[g + \bar{E}_{t-1}[q_{t-1}]\right] \\ &\quad +((1-r)k_t(1) + rk_t(2))\left[q_t - \bar{E}_t[q_t]\right]. \end{aligned} \tag{6.23}$$

The constant terms on the right side must be equal to c_2. The coefficients on q_t are $M(2,1) = (1-r)k_t(1) + rk_t(2)$. Finally, the remaining terms must equal $M(2,2)p_{t-1}$. Substituting for $M(2,1)$ in $M(2,2)$ and rearranging yields $M(2,2) = 1 - (1-r)k_t(1) - rk_t(2)$. The fact that there exist coefficients that solve this system of equations verifies the conjectures.

The final step is to solve for the Kalman gain k_t. To do this, use the Kalman filter formulas (2.9) and (2.10), substituting in the above solution for M. This produces a set of two equations in two unknowns k_t and the posterior variance Σ_t. The solution takes the form $M(2,1) = 1/2[-\gamma + (\gamma^2 + 4\gamma)^{1/2}]$, where $\gamma \equiv (1-r)\sigma_u^2/\sigma_v^2$. Recall that $M(2,1)$ is the sensitivity of the average price to GDP. This sensitivity is decreasing in the degree of pricing complementarity r. That means that more complementarity in prices reduces the sensitivity of prices to innovations in GDP ($M(2,1)$ falls) and increases the extent to which prices depend on the previous period's prices ($M(2,2)$ rises). In other words, the recipe for inertia is a mix of complementarity in actions and heterogeneous information.

6.3 Inattentiveness in Price-Setting

Inattentiveness is a term used to describe models where firms can occasionally observe the entire history of events. It is as if the price-setter knows nothing about demand or monetary innovations until he, one day, receives a newspaper and fully updates his knowledge of the past and present state of the world. In Mankiw and Reis (2002), information arrives randomly, according to a Poisson process. Gabaix and Laibson (2002) and Reis (2006) introduce information choice. Information updates are costly and firms choose when to incur that cost.

Dynamic models with information choice are notoriously hard to solve. Inattentiveness simplifies these problems by making past learning choices irrelevant each time a firm decides to learn. Using a learning technology where acquiring information means learning all past shock realizations perfectly is a way of truncating the state space. Inattentiveness also offers a succinct way of describing agents' information. Each information set is summarized by the last date at which the agent acquired information.

This model, a version of Reis (2006), builds on the quadratic-loss model from the previous chapter.[1] As before, coordination motives in actions (here, price-setting) imply coordination motives in information acquisition. This generates additional inertia, beyond that in Woodford (2002). The additional effect is that, as firms move away from perfect information, the value of any single firm becoming informed diminishes. As firms acquire information less frequently, prices become more sticky.

The model Time is discrete and infinite. There is a measure 1 continuum of firms, indexed by i. Each firm's objective is to minimize their loss function

$$E\left\{\sum_{t=0}^{\infty}\beta^t\left[\left(p_t^i - p_t^*\right)^2 + D_t^i C\right]\right\}, \tag{6.24}$$

where p_t^i denotes firm i's (log-)price in period t; $D_t^i \in \{0,1\}$ is its decision to acquire information (update); p_t^* is an unknown, stochastic target price that a firm with full information would set; $\beta \in (0,1)$ is the firm's discount rate, and $C > 0$ is the cost of information acquisition.

As in New Keynesian models of monopolistic competition, the target price is

$$p_t^* = (1-r)m_t + rp_t \tag{6.25}$$

where m_t is the level of (log-) nominal demand in period t; $m_t - p_t$ is the log of real demand, $p_t = \int_0^1 p_t^i di$ is the average log price; $r > 0$ measures strategic complementarity or real rigidity in price setting. For simplicity, we assume that demand follows a random walk, with innovations $\varepsilon_t \sim i.i.d.N(0,\sigma^2)$

$$m_t = m_{t-1} + \varepsilon_t. \tag{6.26}$$

[1]The quadratic objective function comes from a second-order approximation to a micro-founded model. Ball and Romer (1990) derive this price-setting model from first principles. A technical appendix to Hellwig and Veldkamp (2009) shows that their foundations produce the same objective in a setting with costly information.

Information Choices If firm i last updated in period $\hat{\tau}$, it enters period t with an information set that contains the demand realizations from every period up to and including $\hat{\tau}$: $I_{\hat{\tau}} = \{m_\tau\}_{\tau=0}^{\hat{\tau}}$. If this firm updates in the current period ($D_t^i = 1$), its information set contains all demand realizations up to the current date: $I_{i,t} = I_t = \{m_\tau\}_{\tau=0}^{t}$. If the firm does not update at date t ($D_t^i = 0$), its information set is $I_{\hat{\tau}}$; it does not observe any new information about the state, including endogenous variables like the aggregate price level. The idea is that firms can always see prices, but using them to infer demand or to re-compute their own optimal price incurs a cost.

The following notation describes aggregate information choices. At date t, let $\lambda_{t,\hat{\tau}}$ denote the measure of firms who last updated in period $\hat{\tau} \leq t$, and $\Lambda_{t,\hat{\tau}} \equiv \sum_{\tau=\hat{\tau}}^{t} \lambda_{t,\tau}$ be the measure of firms who have last updated between dates $\hat{\tau}$ and t. Let $D_{t,\hat{\tau}} \in [0,1]$ denote updating choices: the probability that a firm who last updated in period $\hat{\tau}$ will update in period t. $\{\lambda_{t,\hat{\tau}}\}$ is defined recursively as $\lambda_{t,\hat{\tau}} = (1 - D_{t,\hat{\tau}})\,\lambda_{t-1,\hat{\tau}}$ for $\hat{\tau} < t$; the measure of firms who have up-to-date information is $\lambda_{t,t} = \sum_{\tau=0}^{t-1} D_{t,\tau}\lambda_{t-1,\tau}$.

Equilibrium An equilibrium is a sequence of information choices by every firm i $\left\{D_t^i\right\}$, and prices $\{p_{i,t}\}$, that are $I_{i,t}$-measurable and maximize (6.24), taking as given the choices of all other firms.

Prices and Indirect Utility The first-order condition of (6.24) with respect to price dictates that firm i that last updated at date $\hat{\tau}$, sets a price equal to its expected target price at time t: $p_{it} = E\left(p_t^*|I_{\hat{\tau}}\right) = (1-r)E\left(m_t|I_{\hat{\tau}}\right) + rE\left(p_t|I_{\hat{\tau}}\right)$. Since demand is a random walk, $E\left(m_t|I_{\hat{\tau}}\right) = m_{\hat{\tau}}$. We can then guess and verify that price is also a random walk, $E\left(p_t|I_{\hat{\tau}}\right) = p_{\hat{\tau}}$. Thus, the average of all firms' prices is a weighted sum of the expected target price of firms who have last updated their information in period $\hat{\tau}$, $p_t = \sum_{\tau=0}^{t} \lambda_{t,\tau} E\left(p_t^*|I_\tau\right)$ $= \sum_{\tau=0}^{t} \lambda_{t,\tau}\left((1-r)m_\tau + rp_\tau\right)$. Recursively substituting in for p_τ reveals that the average price is a weighted sum of all past demand innovations:

$$p_t = \sum_{\tau=0}^{t} \frac{\Lambda_{t,\tau}(1-r)}{1 - r\Lambda_{t,\tau}} \varepsilon_\tau. \tag{6.27}$$

The proof of this result is left as an exercise.

Substituting this result into (6.25) tells us that the target price process is $p_t^* = \sum_{\tau=0}^{t} \frac{1-r}{1-r\Lambda_{t,\tau}} \varepsilon_\tau$. Firms who last updated at date $\hat{\tau}$ set a price $E\left(p_t^*|I_{\hat{\tau}}\right) = \sum_{\tau=0}^{\hat{\tau}} \frac{1-r}{1-r\Lambda_{t,\tau}} \varepsilon_\tau$. Their expected one-period loss depends on all the demand innovations since the last update:

$$L_{t,\hat{\tau}} = E\left[E\left(p_t^*|I_{\hat{\tau}}\right) - p_t^*|I_{\hat{\tau}}\right]^2 = \sum_{\tau=\hat{\tau}+1}^{t} \left(\frac{1-r}{1-r\Lambda_{t,\tau}}\right)^2 \sigma^2. \tag{6.28}$$

The longer it has been since a firm has updated its information, the higher are its incentives to update in the current period. Consequently, an equilibrium is characterized by threshold dates such that firms who last updated at date $\hat{\tau} < \tau_t^*$ update at date t, while those who last updated at $\hat{\tau} > \tau_t^*$ find not updating strictly optimal.

Complementarity in Information Choice When prices are complements ($r > 0$), there is complementarity in information acquisition (updating). When prices are strategic substitutes ($r < 0$), the converse is true. This general principle in static models (section 4.4) re-appears in dynamic price-setting. Information complementarity is important because it generates delays in price adjustment, which price-setting models are designed to explain.

To see where updating complementarity arises, consider firms' per-period loss from not updating (equation 6.28). For any $\tau = \hat{\tau}+1, ..., t$, $\partial L_{t,\hat{\tau}}/\partial \Lambda_{t,t-\tau} > 0$, if and only if $r > 0$. The more firms are aware of a shock that has occurred since the firm last updated, the higher is the per-period loss of not being aware of this shock, if prices are complementary. This is the complementarity in updating decisions: the more recently other firms have updated, the higher is the cost to a firm of not updating in the current period.

The complementarity in updating delays price adjustment. To illustrate this, we next consider one equilibrium of this updating game. This particular equilibrium has been the focus of previous work (Reis, 2006). In this equilibrium, updating decisions are *staggered*, meaning that all firms update after a fixed number of periods T, and each period a fraction $1/T$ of the firms updates. This means that if $\tau < t - T$, then $\lambda_{t,\tau} = 0$ and $\Lambda_{t,\tau} = 1$, but if $t - T < \tau \leq t$, then $\lambda_{t,\tau} = 1/T$ and $\Lambda_{t,\tau} = (t - \tau + 1)/T$. Therefore, a firm who last updated at date $\hat{\tau}$ has 1-period loss at date t

$$L_{t,\hat{\tau}} = \begin{cases} \sigma^2 \sum_{v=1}^{t-\hat{\tau}} \left(\frac{1-r}{1-rv/T}\right)^2 & \text{if } \hat{\tau} > t-T \\ L_{t,\hat{\tau}+1} + \sigma^2 & \text{if } \hat{\tau} \leq t-T. \end{cases} \tag{6.29}$$

Complementarity ($r > 0$) generates delays in price adjustment through two channels: First, because many other firms have prices based on old information, firms that do update temper their reactions to recent information. This is the effect Woodford (2002) identified and it shows up here in equation (6.27). Second, complementarity reduces the frequency of information acquisition. This effect shows up as $L_{t,\hat{\tau}}$ decreasing in r in equation (6.29). When pricing complementarity causes firms to temper their reactions to new information, the loss incurred from having old information is smaller. As other firms delay updating, this loss falls even more. Firms that update information less frequently have more inertia in their prices.

More generally, complementarity and covariance are mutually reinforcing. With more incomplete information, the covariance of average prices and demand falls. As that covariance falls, demand innovations contain less information about changes in average price. If demand innovations are less useful for coordinating price-setting, the incentive for a firm to update its information about demand diminishes. If firms update less, the covariance of prices and demand falls even further. This feedback is a key feature of the incomplete-information price-setting models that allows them to match the degree of price inertia in the data.

6.4 Rational Inattention Models of Price-Setting

This model, a version of Maćkowiak and Wiederholt (2009b), has similar objectives to the inattentiveness model from the last section. However, it uses a different learning technology. Instead of choosing when to learn and update completely about all past shock realizations, firms are continuously getting noisy information flows about two random variables, an aggregate and a firm-specific variable. The learning choice is how much attention to allocate to each shock. First, we will explore a partial equilibrium model and then discuss how the authors put the mechanism in a general equilibrium model.

The model In each period t, a firm i chooses its price to maximize its expected profit π,

$$\max_{P_{it}} E[\pi(P_{it}, \bar{P}_t, Y_t, Z_{it})|s_i^t]. \tag{6.30}$$

Profit depends on the firm's own price P_{it}, the average of other firms' prices $\bar{P}_t$, an aggregate state (aggregate output) Y_t and a firm-specific shock (productivity) Z_{it}. The expectation of profits is conditioned on the firm's information set, which is the sequence of all signals they have observed up to time t, s_i^t.

The authors solve for the steady-state of the model. Then, they re-express the objective in terms of log-deviations from that steady state. For example, $x = \ln X_t - \ln X^{steady\ state}$ and the profit function is now $\hat{\pi} = \pi(P^{ss}e^{p_{it}}, P^{ss}e^{\bar{p}_t}, Y^{ss}e^{y_t}, Z^{ss}e^{z_{it}})$. Next, they do a second-order Taylor approximation of the objective around the steady state. They show that the optimal price that comes out of the first-order condition is

$$p_{it} = E[\bar{p}_t - \frac{\pi_{PY}}{\pi_{PP}}y_t - \frac{\pi_{PZ}}{\pi_{PP}}z_{it}|s_i^t] \tag{6.31}$$

where $\pi_{ij} \equiv \partial^2\pi/\partial i\partial j$ and $\pi_{PP} < 0$. Notice that the coefficient on the average price in the price-setting rule is positive. This is a coordination motive in price-setting.

The authors rewrite this optimal price by breaking it into two pieces. The first term $\Delta_t \equiv \bar{p}_t - \frac{\pi_{PY}}{\pi_{PP}}y_t$ is the profit-maximizing response to aggregate conditions. The second term is the optimal response to idiosyncratic conditions.

$$p_{it} = E[\Delta_t - \frac{\pi_{PZ}}{\pi_{PP}}z_{it}|s_i^t] \tag{6.32}$$

Finally, the expected loss from not setting the full-information optimal price p_{it}^* takes a quadratic form (because of the second-order Taylor expansion)

$$EL = \frac{|\pi_{PP}|}{2}E\left[(p_{it} - p_{it}^*)^2\right]. \tag{6.33}$$

Firms minimize this expected loss. The aggregate and firm-specific shocks are assumed to be white noise processes, with mean zero and variances σ_Δ^2 and σ_z^2.

Information choice Firms get signals each period about the aggregate and the firm-specific shock. The aggregate and firm-specific signals are not correlated with each other, nor are the signal errors correlated across firms.

$$s_{1it} = \Delta_t + \epsilon_{it} \tag{6.34}$$

$$s_{2it} = z_{it} + \psi_{it} \tag{6.35}$$

where $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$ and $\psi_{it} \sim N(0, \sigma_\psi^2)$. The assumption that signal noise is independent across agents implies that all signals are *private information*. As in chapter 4.6, this ensures a unique equilibrium.

What firms are choosing is the precision of the two signals. This choice is bounded by a constraint on the mutual information of signals and the two random shocks. Recall that the entropy (mutual information) constraint limits the determinant of the posterior variance-covariance matrix of the shocks, relative to their prior variance-covariance matrix (chapter 3.2). Since the z and Δ shocks are independent, their variance-covariance matrix is diagonal and its determinant is the product of the variances. Therefore, the entropy (capacity) constraint can be expressed as

$$\frac{\hat{\sigma}_\Delta^{-2}\hat{\sigma}_z^{-2}}{\sigma_\Delta^{-2}\sigma_z^{-2}} \leq e^{2K}. \tag{6.36}$$

There is an additional constraint that the signal variances cannot be negative. This implies that the posterior variances are always weakly lower than the prior variances.

Each choice of signal precision will map one-for-one into a choice of posterior belief precision about the two shocks: $\hat{\sigma}_\Delta^{-2} = \sigma_\Delta^{-2} + \sigma_\epsilon^{-2}$ and $\hat{\sigma}_z^{-2} = \sigma_z^{-2} + \sigma_\psi^{-2}$. Thus, the information choice problem can be expressed as a choice of posterior variances: Choose $\hat{\sigma}_\Delta^2$ and $\hat{\sigma}_z^2$ to minimize (6.33) subject to (6.36), $\hat{\sigma}_\Delta^2 \leq \sigma_\Delta^2$, and $\hat{\sigma}_z^2 \leq \sigma_z^2$.

Optimal Attention Allocation Substituting the optimal price (6.31) into the objective (6.33), and using the fact that $E[(\Delta_t - E[\Delta_t|s_{1it}, s_{2it}])^2] = \hat{\sigma}_\Delta^2$ and $E[(z_{it} - E[z_{it}|s_{1it}, s_{2it}])^2] = \hat{\sigma}_z^2$, yields

$$EL = \frac{|\pi_{PP}|}{2}\left(\hat{\sigma}_\Delta^2 + \left(\frac{\pi_{PZ}}{\pi_{PP}}\right)^2 \hat{\sigma}_z^2\right). \tag{6.37}$$

Since the expected loss is increasing in both conditional variances, the capacity constraint will always bind. Use the capacity constraint (6.36) to substitute out $\hat{\sigma}_z^{-2}$. The first-order condition of the resulting single-variable optimization is

$$\frac{|\pi_{PP}|}{2}\left(1 - \left(\frac{\pi_{PZ}}{\pi_{PP}}\right)^2 \frac{e^{-2K}\sigma_\Delta^2\sigma_z^2}{\hat{\sigma}_\Delta^4}\right) = 0. \tag{6.38}$$

Under the optimal allocation of attention, the ratio of posterior to prior precision of beliefs about the aggregate shock is

$$\frac{\hat{\sigma}_\Delta^{-2}}{\sigma_\Delta^{-2}} = \left|\frac{\pi_{PP}}{\pi_{PZ}}\right| e^K \frac{\sigma_\Delta}{\sigma_z}. \tag{6.39}$$

This, of course, is only the solution if it is not less than 1 and not greater than e^{2K}. If the ratio were less than 1, the posterior variance would be higher than the prior variance. This means that the firm forgets; it knows less about Δ after observing its signal than it did before. We do not allow firms to choose to forget information in this way. If the ratio were greater than e^{2K}, then the firm would be acquiring more information than its capacity allow. If the ratio is between 1 and e^{2K}, then the corresponding conditional variance of the firm-specific shock is $e^{-2K}\sigma_\Delta^2\sigma_z^2/\hat{\sigma}_\Delta^2$.

Interpreting the solution When the ratio $\hat{\sigma}_\Delta^{-2}/\sigma_\Delta^{-2}$ is one, its minimum, then no attention is being paid to aggregate (Δ) shocks. Greater values mean more attention to aggregates.

The optimal attention allocation depends both on properties of the profit function and the unconditional variances of the aggregate and firm-specific shocks. The more sensitive profits are to the interaction between prices and firm-specific shocks, the higher is π_{PZ}. That makes devoting more attention to aggregate shocks less valuable, which raises the optimal posterior variance on aggregate shocks $\hat{\sigma}_\Delta^2$. Conversely, it makes devoting more attention to firm-specific shocks more valuable and lowers $\hat{\sigma}_z^2$. That is not surprising, but helps explain what determines the attention allocation. When aggregate shocks have higher unconditional volatility, it is also optimal to devote more attention to them. Thus, firms want to devote more attention to the shock that their price is more sensitive to and the shock that is more volatile.

The question of how much inertia in price-setting this model generates is ultimately an empirical one. Thus the authors calibrate the stochastic processes to measures of the volatility of nominal aggregate demand shocks and the volatility of firm-level shocks as measured by the average size of a firm's price change. They calibrate the preference parameters to values for price sensitivities used in the monetary literature. They find that firms optimally allocate 94% of their attention to firm-specific conditions. The primary reason that little attention is paid to aggregate shocks is that firm-specific shocks are about 10 times more volatile.

Because price-setters devote so little attention to aggregate shocks, prices do not covary highly with these shocks. As a result, monetary policy has very little contemporaneous effect on prices. Its effect works its way slowly through prices as new information is processed at a slow rate.

The more general lesson is that in environments with multiple random shocks to optimal actions, firms will tend to devote more attention to the most volatile shocks. That can make actions appear unresponsive to the more subtle changes in their environments.

Attention Allocation in a Production Economy The reason monetary economists want to understand what makes prices unresponsive to aggregate monetary shocks is because it is precisely that unresponsiveness that causes monetary policy to have real effects on output and consumption. To gauge those effects, Maćkowiak and Wiederholt (2009a) extend the framework described above from a partial equilibrium with exogenous aggregate output, to general equilibrium. Households choose consumption and wages and firms choose labor inputs and set prices. Each observes only noisy signals about aggregate technology shocks, monetary policy (interest rate) shocks, firm-specific productivity shocks. Firms pay least attention to monetary shocks, pay more attention to productivity shocks and pay the most attention to the micro-level shocks. As before, this result follows from the calibrated relative volatilities of each type of shock. The firm-specific and individual-specific shocks are the most volatile, followed by aggregate productivity and lastly monetary policy.

Modeling rational inattention in an equilibrium model raises lots of questions. For example, if output is unknown, then the agent cannot choose both consumption and savings. They can choose one and the other must be a residual. Another issue is how information revealed by equilibrium prices is incorporated into the information constraint. Here, instead of writing the information flow constraint in terms of the prior and posterior variances of each shock, Maćkowiak and Wiederholt constrain the amount of information that actions contain about the optimal, full-information action. For example, suppose a firm chooses price p and labor input l. Given full information about the state of the economy, the firm would choose p^*, l^*. Then, this constraint would bound the mutual information $\mathcal{I}(\{p, l\}, \{p^*, l^*\})$. This constraint represents the same physical learning process as the constraint on signal precisions. If actions are conditioned on any kind of signal that contains information about the true state, mutual information $\mathcal{I}$ rises. Thus, the constraint incorporates the effect of all information from both exogenous signals and endogenous sources such as market prices. Because learning from market prices requires capacity, this model does not need to introduce shocks that make market prices noisy to keep information heterogenous. The information flow restriction ensures that all firms extract noisy heterogeneous information, even from public signals like prices.

The key result is that changes in monetary policy produce delayed, hump-shaped consumption and output responses, just as they do in the data. While Maćkowiak and Wiederholt (2009a) solve pieces of this model analytically, the main results are numerical in nature.

I refer the reader to the paper for more details.

6.5 Are Prices State-Dependent or Time-Dependent?

One way of seeing the contribution of information choice models in monetary economics is that they advance the debate about whether prices adjust in response to changes in economic conditions (state-dependence) or whether they adjust at regular intervals (time-dependence). An example of a state-dependent price adjustment model is one with a fixed adjustment cost or "menu cost" to update prices. In such a model, whether the firm decides to update its price depends on its state variables, specifically, the distance between its current price and its optimal prices. An example of a time-dependent price adjustment model is one where firms update prices every T periods. This is often referred to as Calvo (1983) pricing.

The problem with state-dependent pricing is that it generates little rigidity in aggregate prices. Golosov and Lucas (2007) show that even though very few firms adjust prices each period, the aggregate price adjusts quickly to changes in the state because it is the firms whose prices have the farthest to adjust that do the adjusting. This *selection effect* is mitigated in time-dependent models because the firms do not choose to adjust when adjustment is most valuable. They adjust when their time comes.

However, many oppose Calvo pricing because it lacks microeconomic foundations. Why would a firm wait the required number of periods to adjust its price if some event just happened that makes non-adjustment very costly? Models with inattentiveness, like Reis (2006), attempt to provide micro-foundations for a Calvo-like model. Reis' answer to the criticism of the Calvo model is that when an event that warrants a big price adjustment takes place, firms may not react because they are not paying attention. They are not paying attention to market events because attention is costly.

Models with rational inattention, like Woodford (2002) and Maćkowiak and Wiederholt (2009b) have a constant flow of information that firms can react to. Thus, they are more like state-dependent models than time-dependent models. But imperfect information reduces the dependence of prices on the state. These models generate non-trivial price rigidity because the imperfect information works to mitigate the selection effect. If firms do not know how far their prices are from the optimal price, then some firms who should adjust do not adjust and some of the price-adjusters make only small adjustments. (See also Moscarini (2004) on this point.)

Some of the most recent work on information frictions in monetary economics attempts to integrate both state and time-dependence in one framework. In Woodford (2008), firms face a fixed cost of acquiring perfect information and updating their prices, just like the models of inattentiveness. Whereas inattentive firms get no information about the state in between price updates, these firms get new noisy signals every period, just like the firms with rational inattention. The state-dependent and time-dependent models are two limiting cases of this model. When the signal observed each period is perfectly precise, then this becomes a simple menu-cost model with a fixed cost of price adjustment. That is state-dependence. When the signal contains no information, then the firm's decision about whether to update its price depends only on the number of periods since it last updated. That is time-dependence. Calibrating such a model could give us an idea of which assumption is closer to the truth.

A Hybrid Model of Rational Inattention and Inattentiveness Woodford (2008) constructs a model where a firm chooses how much information to acquire, rather than how to allocate its attention. Like Reis (2006), the firm can pay a fixed cost at discrete dates to observe perfect information about the current and past states of the economy. But in between these updates, it can observe a flow of noisy information about the state, as in Maćkowiak and Wiederholt (2009b). This is meant to capture the idea that it is costly both to adjust prices and to monitor the economy to know when to adjust.

Woodford writes the problem in terms of the *price gap* x, the discrepancy between a firm's chosen price and the price that would be optimal under full-information. The firm's prior belief is that $x \sim f(x)$. At each date, the firm observes a signal s about x. The mutual information of s and x is denoted $I(s,x)$. The firm can choose the precision of the signal s. A more precise signal will have higher mutual information I, which has a constant marginal cost θ. This is the *rational inattention* part of the model.

The firm uses its signal s to update its beliefs about x using Bayes' Law. Then, it makes a choice about whether or not to adjust prices. The adjustment choice variable is $\delta(s)$ which is 1 if prices adjust and 0 otherwise. If the firm adjusts its price, it gets perfect information about the current (and past) values of x. This is the *inattentiveness* part of the model.

The objective of the firm depends on the benefit from updating, net of the updating cost $L(x)$. This benefit is the firm's value if it pays the updating cost, updates, and sets the price gap to zero, minus its value with price gap x. Thus, the firm chooses a quantity $I(s,x)$ and the function $\delta(s)$ to maximize its objective

$$E[\delta(s)L(x)] - \theta I(s,x). \tag{6.40}$$

The first term is a product because the firm only gets benefit $L(x)$ if it updates and sets $\delta(s) = 1$.

Woodford proves that the optimal policy involves acquiring a signal that takes on only two values ($s \in \{0,1\}$) and using an updating function that prescribes a review whenever the signal is 1 ($\delta(s) = s$). Given these features, the question for the firm is how precise a signal do they want? The precision of the signal is described by a function $\Lambda(x)$ that gives the probability that $s = 1$, given that the true state is x. Woodford calls this a *hazard function* because it is the probability that a firm will decide to fully update its information and reset its price to the optimal full-information price.

To solve for the optimal hazard function, we need to express the mutual information I in terms of $\Lambda(x)$. To express mutual information, we first need to know the probability of state x conditional on signal s. This conditional (posterior) probability is given by Bayes' Law

$$f(x|s=0) = \frac{f(x)(1-\Lambda(x))}{1-\bar{\Lambda}} \qquad f(x|s=1) = \frac{f(x)\Lambda(x)}{\bar{\Lambda}} \tag{6.41}$$

where $\bar{\Lambda}$ is the prior probability of observing $s = 1$: $\bar{\Lambda} \equiv \int \Lambda(x)f(x)dx$.

Recall from chapter 3.2 that the entropy of a variable $x \sim f(x)$ is $H(x) = -E[\ln f(x)]$ and conditional entropy is $H(x|y) = -E[\ln f(x|y)]$. Therefore, the entropy of x, conditional on the signal s is $H(x|s) = -\bar{\Lambda}E[\ln f(x|s=1)] - (1-\bar{\Lambda})E[\ln f(x|s=0)]$. Finally, mutual information is entropy minus conditional entropy (equation 3.1). Replacing the expectation $E[\ln f(x|y)]$ with the integral $\int \ln(f(x))f(x)dx$ and doing the same substitution for the conditional entropy term yields

$$\begin{aligned} I(s,x) \quad = \quad & -\int \ln(f(x))f(x)dx + \bar{\Lambda}\int \frac{f(x)\Lambda(x)}{\bar{\Lambda}} \ln\left(\frac{f(x)\Lambda(x)}{\bar{\Lambda}}\right) f(x)dx \\ & +(1-\bar{\Lambda})\int \frac{f(x)(1-\Lambda(x))}{1-\bar{\Lambda}} \ln\left(\frac{f(x)(1-\Lambda(x))}{1-\bar{\Lambda}}\right) f(x)dx. \end{aligned} \tag{6.42}$$

We can write the log of a product as a sum of logs and thereby extract all the $\ln(f(x))$ terms. They all cancel each other out. Collecting the remaining terms under one integral reveals that mutual information is

$$I(s,x) = \int \left[\Lambda(x)\ln\left(\frac{\Lambda(x)}{\bar{\Lambda}}\right) + (1-\Lambda(x))\ln\left(\frac{1-\Lambda(x)}{1-\bar{\Lambda}}\right)\right] f(x)dx.$$

Next, break up both log ratio terms into the difference of logs. Also, write the integral over four additive terms as the sum of four integrals. Note that $-\ln(\bar{\Lambda})$ and $-\ln(1-\bar{\Lambda})$ do not vary in x and can therefore be extracted from the integral, leaving $-\ln(\bar{\Lambda})\int\Lambda(x)f(x)dx = -\ln(\bar{\Lambda})\bar{\Lambda}$ and $-\ln(1-\bar{\Lambda})\int(1-\Lambda(x))f(x)dx = -\ln(1-\bar{\Lambda})(1-\bar{\Lambda})$ as two of the four terms. Thus,

$$I(s,x) = \int\Lambda(x)\ln(\Lambda(x))f(x)dx + \int(1-\Lambda(x))\ln(1-\Lambda(x))f(x)dx - \bar{\Lambda}\ln(\bar{\Lambda}) - (1-\bar{\Lambda})\ln(1-\bar{\Lambda}).$$

Combine the first two terms in one integral and examine the term inside that integral. Using the definition of entropy again, note that this term is the entropy of the signal s, conditional on the hazard function $\Lambda(x)$. $E[log(p(s=0|\Lambda(x))) + log(p(s=1|\Lambda(x)))] = -H(s|\Lambda(x))$. Likewise, the last two terms are the entropy of s, conditional on $\bar{\Lambda}$, the prior probability: $H(s|\bar{\Lambda})$.

Substituting mutual information into the firm's objective function (6.40) and using the fact that $\Lambda(x)$ is the probability $Prob[\delta(s)=1|x]$, delivers the firm's maximization problem

$$\max_{\Lambda(x)} \int \Lambda(x)L(x) - \theta H(s|\Lambda(x))f(x)dx + \theta H(s|\bar{\Lambda}). \tag{6.43}$$

Woodford establishes that a first-order approach characterizes the unique global solution to this problem. In each state x that has positive measure ($f(x) > 0$), the first derivative of the objective with respect to Λ must be zero:

$$\left(L(x) - \theta\frac{\partial H(s|\Lambda(x))}{\partial\Lambda(x)}\right)f(x) + \theta\frac{\partial H(s|\bar{\Lambda})}{\partial\bar{\Lambda}}\frac{\partial\bar{\Lambda}}{\partial\Lambda(x)} = 0 \qquad \forall x \text{ such that } f(x) > 0 \tag{6.44}$$

Simple differentiation reveals that $\partial H(s|\Lambda(x))/\partial\Lambda(x) = \ln(\Lambda(x)/(1-\Lambda(x)))$ and $\partial\bar{\Lambda}/\partial\Lambda(x) = f(x)$. Since we are looking at states where $f(x) > 0$, we can cancel the $f(x)$ terms and express the optimal hazard function as

$$\ln\left(\frac{\Lambda(x)}{1-\Lambda(x)}\right) = \ln\left(\frac{\bar{\Lambda}}{1-\bar{\Lambda}}\right) + \frac{L(x)}{\theta}. \tag{6.45}$$

This condition is what pins down the optimal degree of rational inattention. One of the features of this model is that there is little incentive to allocate much attention to x because adjusting the price allows the firm to get periodic updates with perfect information.

Woodford goes on to describe how in a dynamic model, prior beliefs $f(x)$ arise endogenously from past signals. He also links the payoff function $L(x)$ to economic fundamentals in an equilibrium model of monopolistically competitive firms.

A calibrated version of this model delivers a realistic degree of price inertia to changes in monetary policy. The reason is that some firms whose prices are far away from the optimal prices (a large $|x|$) do not know this. Because of the information frictions, such firms have prices that are "sticky," meaning that they fail to react to changes in the monetary environment.

6.6 Broader Themes and Paths for Future Research

The mechanisms explored in this chapter, particularly the interaction between coordination motives and heterogeneous information, can be used to explain many situations where aggregate actions exhibit inertia.

Consumption Inertia and Excess Sensitivity Two empirical puzzles about consumer behavior cannot be explained by a standard, rational expectations, permanent income hypothesis (PIH) model. The first puzzle is excess consumption smoothness, whereby aggregate consumption growth is much smoother than aggregate income growth in the US data than the one predicted in line with PIH. The second is the excess sensitivity puzzle, whereby aggregate consumption is actually more sensitive to anticipated income innovations than one predicts based on PIH.

Luo (2008) shows that a model with rational inattention can explain these facts. Agents have quadratic utility and face permanent (random walk) and transitory (moving average) shocks to their income. They do not observe income, but can acquire noisy signals about it.

The main result is that, a smaller information processing capacity dampens the response of consumption to the permanent and transitory income shocks in the short run, but amplifies the response in the long run. The short-run response is dampened because, as we have seen many times now, imperfect information makes actions less correlated with the unobserved state variable. The long-run response is amplified because an agent who has high income but does not increase their consumption accumulates wealth that allows them to have even higher consumption in future periods.

Insensitivity of Asset Prices to Some Shocks Just like goods prices, equity prices can be slow to incorporate some kinds of news and quick to incorporate others. Hong et al. (2007a) give the example of Amazon's stock price. It initially reacted mostly to information about numbers of clicks on its website and not to earnings announcements, even though earnings announcements seemed more relevant to future dividends. In the savings-consumption context, tax cuts seem to affect consumption behavior differently from equivalent tax rebates. In the world of corporate debt markets, investors react more to a ratings downgrade than to an upgrade. These types of phenomena could potentially be explained by a model of attention allocation. Maćkowiak and Wiederholt (2009b) show us that more volatile variables are more valuable to pay attention to. This same effect could take the form of a dynamic reallocation effect: In times when a variable has higher expected volatility, firms track it closely and react quickly to it. When a variable is expected to be relatively constant, changes in its value may not affect firms' behavior because they choose to allocate their attention to something else.

Similarly, there are events where private information does not effect aggregate actions but public information does. For example, some news stories do not affect the price of a stock the first time they are printed, but do when they are later reprinted in a higher-circulation publication (Huberman and Regev, 2001). An effect that Reis (2006) highlights is that when firms want to coordinate, the incentive to acquire information that is closer to common knowledge is stronger than the incentive to acquire information that most others will choose not to observe.

Overreaction to Recent News in Asset Markets In chapter 6.2, we saw how the combination of coordination motives and heterogeneous information was a potent force for inertia. The idea was that if you wanted to take an action (e.g., set a price) that is close to the actions others choose, you want to react more to public information than to private information. When firms get updated information only infrequently, older information is more public. Newer information is known by only the firms that have updated their information sets since the new event took place. Thus, firms put more weight on old, more public information which makes their actions slow to evolve.

But what if the strategic motives in actions are characterized by substitutability, rather than complementarity? Chapter 7 will show that investment decisions are characterized by such substitutability. If firms want to take actions that differ from what others do, they want to weight private information more and public information less. When firms get updated information only infrequently, newer information is more private. Thus, weighting newer information more is optimal. Such logic might explain why there is excess volatility in asset markets.

Ratings Agencies The tools in this chapter could also be used to ask normative questions. For example, a ratings firm has a large number of assets to follow and has limited information processing resources available to re-rate assets. Which assets should they collect more information on to potentially re-rate? If ratings agencies follow this optimal procedure for re-rating, can firms take actions that manipulate this process? Can they make investments that minimize their chances of being re-rated when their performance is poor and maximize it when performance is strong? A framework like the one presented here where firms re-set ratings, instead of prices could be useful for answering such questions.

Monitoring and Moral Hazard Another application of rational inattention and inattentiveness models is to re-think questions about optimal monitoring that arise in applied micro and corporate finance. For example, if a manager has to monitor many different operations, how does he allocate his attention between them? How does the managers compensation scheme affect his allocation choice?

Central Bank Transparency Revisited The main question raised in the previous chapter is whether a central bank or other public authority should reveal information publicly. Morris and Shin (2002) argue that if coordination produces private benefits, but not public benefits, then too much coordination, which is facilitated with public information, can be welfare-reducing. In Amador and Weill (2009b), when agents know lots of public information, their actions reveal little private information and others learn little new from observing them.

This chapter introduced information choice not by the government, but by the firms who set the prices. The possibility that the government's information disclosure might interact with firms' information choice could create another rationale for withholding public information. The argument is that public information can deter information acquisition. If agents with public information acquire less private information, aggregate information could suffer and social welfare with it.

6.7 Exercises

1. Prove the result in equation (6.27). Conjecture that the price p_t is a linear combination of past money supply innovations $\epsilon_{t-\tau}$. Then, use equations (6.25) and (6.26) to verify the conjecture and solve for the undetermined coefficients.

2. Suppose that all the firms updated their information periods ago but no firm has updated it since. Is it an equilibrium? Express the current price as a function of demand innovations.

3. If more firms planned more recently, how does this affect one firm's loss from not planning? Hint: express your answer as a partial derivative.

4. Prove that if $r > 0$ and if all firms choose to update in even periods, then any given firm strictly prefers to update in even periods rather than in the odd periods.

5. Verify the expression for the optimal price (equation 6.31) in Maćkowiak and Wiederholt (2009).

6. In Maćkowiak and Wiederholt (2009) assume now that the firms process information subject to a linear information processing constraint. What is the solution to this version of the model? Explain it carefully.