Monetary policy for rationally inattentive economies with staggered price setting

Fang Zhang

California State University-Fullerton, 800 N. State College Blvd., Fullerton, CA 92620, United States

Abstract

The paper examines the optimal monetary policy when firms are constrained by information processing capability and infrequent price adjustments. Firms' information processing limit gives rise to imperfect knowledge about macroeconomic aggregates and endogenous information learning contingent on the monetary policy. Staggered price setting introduces the observed price duration and additional policy tradeoffs resulting from the interactions between nominal rigidities and imperfect information processing. The integrated model implies an optimal policy that commits to complete price stabilization in response to natural rate shocks but not in response to markup shocks. In the presence of markup shocks, it is optimal for the central bank to focus on price stabilization in the initial periods following markup shocks and shifts the emphasis to output gap stabilization later. Moreover, larger information capacity, stronger complementarities and more persistent shocks require more aggressive price stabilization in the short-run.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

This paper studies the optimal monetary policy in a Taylor staggered-pricing economy in which firms pay limited attention to aggregate variables.

What policy should a central bank follow? It is a classic question in macroeconomics. The answers reflect economists' belief about how the economy works, what should be valued, what monetary policy is feasible, and hence what explicit assumptions and models should be used.

To account for the real effect of monetary policy, the literature mainly focuses on two types for models for monetary policy analysis: imperfect information models and models with short-run nominal frictions. The widely used New Keynesian models, such as Goodfriend and King (1997), and Clarida et al. (1999), assert that firms' price adjustment is infrequent and time-dependent. Despite their ability to generate monetary neutrality and appealing microfoundations, it is found that such models encounter difficulties in generating plausible effects of monetary policy, such as counterfactual costless deflations (Ball, 1994) and failure to generate a delayed and persistent effect on inflation (Mankiw and Reis, 2002). The basic New Keynesian model thus has been modified by incorporating a backward-looking component into the Phillips curve (e.g. Gali and Gertler, 1999). However, the modification is ad hoc, and it has been criticized for not able to explain the...
lack of inertia in early U.S. monetary regime when the economy operated under a gold standard and for ignoring the Lucas critique similar to the conventional Phillips Curve (Ball et al., 2005).

Another part of the literature focuses on the information structure in monetary models. Slowly disseminating or incomplete information can produce the observed sluggish inflation response to a monetary shock and account for monetary nonneutrality. For example, Lucas (1972) assumes that producers solve a signal extraction problem to infer aggregate prices. However, this idea has been criticized since macroeconomic data is widely available, and such information should be considered by a rational person. Mankiw and Reis (2002) sticky information model assumes frictions in updating information rather than updating prices. It successfully generates the inflation-output tradeoff and monetary nonneutrality, but suffers similar problems as Lucas (1972). The main mechanism for generating inertial inflation behavior is firms’ price setting decision based on outdated information, but it is puzzling why some firms would use very old information when most macroeconomic variables are publicly available with little delay.

Rational inattention models, as in Sims (2003), may be able to solve this dilemma. The model claims that agents’ inability to process, not obtain, the data perfectly prevents them from obtaining full information. Economic agents have a limited capacity to process information “attention”, and they rationally choose to be less attentive to some available information and fail to incorporate full information in their decision-making. An information processing constraint is an appealing way to model information frictions for monetary models for several reasons. The method, discussed in Sims (2003) and initiated in information theory, is able to approximate information processing procedure while stepping away from the psychological details. Furthermore, optimality with limited information processing resources leads to information endogeneity. Agents’ information learning and expectation are policy contingent, making the model robust to Lucas critique.

This paper makes progress towards determining the optimal monetary policy by integrating staggered pricing models with an information friction resulting from limited information processing capability. Firms have multi-period overlapping price contracts as in Taylor (1979), and they optimize their information processing to obtain the most accurate information given their constraints. Both information processing limits and infrequent nominal adjustment are empirically relevant features. The integrated model is capable of generating a delayed price response to shocks and an amplified output response to monetary disturbances.

There are two types of shocks that cause aggregate fluctuations in the model: natural rate shocks, which are driven by technology innovations and change the efficient level of output, and markup shocks, which cause fluctuations without affecting the efficient level of output. The aggregate fluctuations lead to two inefficiencies. First, aggregate output may deviate from the efficient output as the shocks disturb firms’ pricing and production process (real effect of shocks). Second, price dispersion arises, which leads to inefficient allocation of the production resources. In the model, the monetary authority solves a Ramsey problem and chooses a money supply rule to minimize these two inefficiencies.

The effectiveness of monetary policy depends on how the nominal and information frictions affect firms’ pricing decisions. Monetary policy affect output gap variation through three channels. First, fundamental shocks directly lead to changes in output gap, the magnitude of which can be dampened by policy response. With incomplete information and price adjustment, a change in the money supply leads to a partial adjustment in output gap. Second, staggered pricing introduces an intertemporal link in firms’ prices, allowing for any policy response to have a persistent effect on the output gap. Third, variations in economic fundamentals raise uncertainty. Tracking errors arise as firms’ information processing is imperfect. A policy that facilitates learning improves the quality of information received by firms and leads to more efficient pricing and production decisions. Furthermore, inefficient price dispersion can also be traced to the two frictions in the model. Staggered price contracts lead to price difference across firms depending on their time of adjustment. Information heterogeneity also contributes to price dispersion as firms base their pricing decision on their own information set. A monetary policy that emphasizes price stabilization and facilitates information processing thus can minimize the inefficiency caused price dispersion.

I find that an optimal policy implied by this integrated model should fully accommodate natural rate shocks and stabilize prices. However, when markup shocks are present, there are conflicting policy effects. Offsetting the markup shocks leads to price stabilization, which facilitates learning, reduces relative price distortion, and eliminates persistent shock effects. Intuitively, price stabilization allows for better knowledge of the aggregate price overall and more efficient pricing. It also helps firms that lack the ability to update their prices when a markup shock occurs, since the optimal reset price remains unchanged. Therefore, a price stabilization objective serves the tasks of (1) minimizing inefficient price dispersion and (2) minimizing the inefficient output gap variation due to tracking errors and intertemporal transmission of shock effects through staggered prices. However, markup shocks lead to inefficient fluctuations in output, which requires the central bank to accommodate the shock and let prices fluctuate. Therefore, a careful balance between the effects is necessary, and complete price stabilization is no longer optimal.

The tradeoff between the objectives is determined by firms’ information processing capacity, strategic complementarities in pricing, and shock persistence. I find that it is optimal for the monetary authority to focus more on offsetting markup shocks and stabilizing prices in the short-run, and shifts to accommodating the shock and stabilizing output in the long-run.\footnote{See, for example, Bils and Klenow (2004) and Nakamura and Steinsson (2008) for evidence of infrequent price adjustment and estimates of adjustment frequency; Radner (1992) regarding firms’ limited managerial resources towards collecting and processing information and taking the corresponding actions; Coibion and Gorodnichenko (2012) for evidence of information rigidity and imperfect information modeling favoring rational inattention.}
Larger information capacity, stronger complementarities, and higher shock persistence all lead to more aggressive price stabilization in the short-run. Larger attention also leads to a more rapid shift of policy emphasis to stabilize output in later periods.

The paper is mainly related to three strands of literature. It is related to the rational inattention literature, such as Sims (2003), Mackowiak and Wiederholt (2009), and Woodford (2009). This paper models information structure following the line of this literature, but none of the above studies consider monetary policy problems. My paper also relates to the literature that studies the monetary policy using time-contingent pricing models, such as Fukunaga (2007) and Wolman (2001). Fukunaga (2007) considers a similar Taylor economy, but it studies the effect of an exogenous monetary policy when agents have higher order beliefs. Wolman (2001) studies the optimal monetary policy a Taylor staggered price setting, but does not consider the role of imperfect information. More closely, my paper is related to the literature that considers the policy implications of imperfect information, such as Ball et al. (2005), Adam (2007), and Paciello and Wiederholt (2013). Ball et al. (2005) study optimal monetary policy in a sticky-information setting and predicts price level targeting to delivery the best outcome. Adam (2007) and Paciello and Wiederholt (2013) analyze the optimal monetary policy in a flexible-price economy when agents have limited information processing capacities. A standard rational inattention model with flexible prices has difficulty explaining why prices are fixed for some time (Mackowiak and Wiederholt, 2009). The flexible price assumption, without the aid of an additional mechanism such as higher order beliefs or persistent shocks, also gives rise to implausible prediction to inflation dynamics, such as an immediate and transient response, though the magnitude is dampened. My paper considers Taylor staggered pricing to incorporate the observed fact that firms do not adjust prices frequently. Not only does incorporating nominal rigidities account for the observed price duration and delayed response, it also introduces additional concerns in monetary policy implementation and policy effectiveness when there are interactions between nominal rigidities and rational inattention.

The rest of the paper is organized as follows. Section 2 presents the baseline model and derives key equations regarding firms’ policy-contingent pricing and information processing. Section 3 discusses analytically the effect and tradeoffs in monetary policy implementation, and derives the optimal monetary policy in the simple baseline model. Section 4 extends the analysis to consider a more general policy rule and shock dynamics. Section 5 concludes.

2. The basic model

This section presents the basic model for optimal monetary policy analysis. The economy is populated by monopolistic competitive firms, households, and the government.

2.1. Household

A measure 1 of identical households exist, from which I study the problem of a representative household. The representative household supplies labor, consumes goods produced by the firms, obtains utility from consumption and leisure, and demands government-supplied money in an amount equal to their nominal consumption spending.

The preference of the representative household is given by

\[ U(C_t, L_t) = \frac{C_t^{1-\sigma} - 1}{1-\sigma} - \frac{L_t^{1+\psi}}{1+\psi}. \]

(1)

where \( C_t \) is an index of consumption goods, defined as \( C_t = \int_0^1 (C_{it})^{(\theta - 1)/\theta} di^{\theta/(\theta - 1)} \), and \( L_t \) is the labor supply in period \( t \). Parameter \( \sigma > 0 \) measures the degree of risk aversion, \( \psi > 0 \) measures the disutility of labor supply, \( \theta > 1 \) measures the elasticity of substitution between different goods.

In every period, the household chooses consumption, labor supply, money balances \( M_t \) and nominal bond holdings \( B_t \) to maximize its utility, subject to its budget constraint

\[ \int_0^1 P_t C_{it} di + M_t + B_t \leq W_t L_t + N_t + T_t + M_{t-1} + R_{t-1} B_{t-1} \]

(2)

by taking as given the nominal wage rate \( W_t \), nominal aggregate profit \( N_t \), lump-sum tax transfers \( T_t \), and the prices of all consumption goods \( \{P_{it}\}_{i=0}^1 \). \( R_{t-1} \) is the nominal return on bonds between \( t-1 \) and \( t \).

As is well-known (e.g. Woodford, 2003), utility maximization with this form of utility function implies that the demand for each good depends on aggregate spending and the good’s relative price

\[ C_{it} = \left( \frac{P_{it}}{P_t} \right)^{-\theta} C_t. \]

(3)

\[ ^2 \] Bils and Klenow (2004) find that the average price duration is 7 months; Nakamura and Steinsson (2008) find the average duration to be between 8 and 11 months by using a more detailed data set.
$P_t$ is the price index for the aggregate consumption good $C_t$, defined as $P_t = \left[ \int_0^1 p_t^1 \exp(-\theta \tau) \, d\tau \right]^{1/(1-\theta)}$. In log form, the demand function for good $i$ is given by

$$C_{it} = c_i - \theta(p_{it} - p_t).$$  \hspace{1cm} (4)

The household also faces a cash in advance constraint: $p_t + c_t \leq m_t$, where the lower case variables are in their log form.

### 2.2. Firms

There is a continuum of firms indexed by $i$. Firms use labor as an input, sell their goods in a monopolistically competitive market, and face a constant elasticity of demand. Firms maximize real profit by choosing prices, production, and their information structure. Firms are subject to both nominal and information frictions: they may only adjust their prices every two periods; they are endowed with only a finite information processing capability, which limits their ability to acquire information about market conditions.

Firm $i$’s production technology is linear in labor

$$Y_{it} = A_i L_{it},$$ \hspace{1cm} (5)

where $L_{it}$ is the labor used by firm $i$, $Y_{it}$ is firm $i$’s output, and $A_i$ is an exogenous aggregate productivity shifter in period $t$. $\ln(A) \sim N(0, \sigma^2)$. Market clearing requires $C_{it} = Y_{it}$, $C_t = Y_t$, and $L_t = \int_0^1 L_{it} \, d\tau$.

Firm $i$’s real profit is given by

$$(1 - r_t)\hat{P}_{it} Y_{it} - W_t L_{it},$$ \hspace{1cm} (6)

where $\hat{P}_t = p_t / P$ and $W = W / P$ are its relative price and the real wage, and $r_t$ is a proportional sales tax collected by the government, the distribution of which satisfies $\ln(1 - r_t) \sim i.i.d. N(\ln(1 - \tau), \sigma^2)$. **2.2.1. Full information benchmark**

I first consider firms’ optimization problem when they have access to full information, which is a special case of the baseline model when firms information capacity is infinity. The derived pricing behavior in this full information case serves as a benchmark for later analysis.

**Appendix A** shows that the price (in log form) that maximizes firms’ real profit in period $t$ takes the form

$$p_t^* = p_t + \alpha\left(1 - \frac{\ln Y_N}{\ln A_t}\right)_t + u_t,$$ \hspace{1cm} (7)

where $\alpha = (\sigma + \psi) / (1 + \theta\psi)$, $y_N$ is the natural level of output, and $u$ reflects random variation in taxes. Parameter $\alpha$ measures the degree of strategic complementarity in firms’ prices. A higher $\alpha$ results in a smaller price adjustment in response to shocks. Too see this, combine Eq. (7), the cash-in-advance constraint, and the market equilibrium condition:

$$p_t^* = (1 - \alpha)p_t + \alpha(m_t - y_N) + u_t.$$ \hspace{1cm} (8)

The natural level of output $y_N$ is defined as the efficient log output level when all goods are priced optimally without any nominal or information restrictions, i.e. $p_t = p^*$, $\forall t$, and the sales tax rate is at its average level $\tau$. It is given by

$$y_N = \frac{1}{\psi + \sigma}\left[\ln A_t - \ln\left(\frac{\theta}{\theta - 1}\right)\right] - \ln(1 - \tau).$$ \hspace{1cm} (9)

I consider $y_N$ to be natural rate shocks as they capture the efficient variations in the output level. Note that $y_N$ is driven by productivity variations. Specifically, when $\tau$ is at a level that corrects the distortion arising from monopolistic competition, i.e. $\tau = 1/\theta$, the natural rate shock is proportional to aggregate productivity shocks, i.e. $y_N^N = (1 + \psi)/(\psi + \sigma)\ln A_t$.

Another fundamental shock in the model is a markup shock $u_t$, given by

$$u_t = \frac{1}{1 + \theta\psi}\ln\left(\frac{1 - \tau}{1 - \tau_t}\right).$$ \hspace{1cm} (10)

It captures the variation in firms’ markup, and can change optimal prices and output without changing the natural level of output, causing inefficient distortions. Both natural rate shocks $y_N^N$ and markup shocks $u_t$ are Gaussian white noise.

The pricing structure follows Taylor (1979). Price contracts have a two-period duration and are staggered: for each period, half the firms write new contracts with new prices and the rest sell their goods based on the old contracts with prices updated last period. I assume that for any period $t$, firm $i \in \left(0, \frac{1}{2}\right]$ can adjust its price in period $t$, and firm $i \in \left(\frac{1}{2}, 1\right]$ uses...

---

5 As long as $\alpha \leq 1$, prices are strategic complements, i.e. firms’ optimal prices change less sensitively than the average price level. Throughout the paper, it is assumed that prices are strategic complements as it is the case of greatest economic interest. This is the case when households are not too risk averse.

4 A markup shock is a standard feature of monetary models, and there are many ways to justify them, such as shifts in firms’ market power, variations in consumer preference, and shifts in the wage bargaining power in the labor market. I model markup shocks with variable taxation following Ball et al. (2005) since it is a simple way to generate time-varying markups. The precise source for the markup shocks is not relevant for the main result of the paper.

Section 4 considers persistent shocks.
the price chosen in period \( t - 1 \). The (log) aggregate price for period \( t \) is given by

\[
p_t = \int_0^{1/2} p_{it}^0 \, dt + \int_{1/2}^1 p_{it}^1 \, dt.
\]

(11)

where \( p_{it} \) is the log price charged by firm \( i \). A superscript 0 indicates a newly updated price, and 1 indicates a price that is one-period old.

For simplicity, I assume that firms’ time discount factor is 1. Firms that can update their prices at period \( t \) optimally set their prices to be the average of the profit-maximizing price over the two periods. I use \( p_t^* \) to denote the optimal reset price for firms that can adjust price in period \( t \) under full information.

\[
p_t^* = \frac{1}{2} p_t^* + \frac{1}{2} p_{t+1}^*|X_t|,
\]

(12)

where \( p_{t+1}^*|X_t \) denotes the rational expectation of period \( t + 1 \)’s profit-maximizing price \( p_{t+1}^* \) using all the information available at period \( t \).

With perfect knowledge about the optimal reset price \( p_t^* \), firms that can update prices set their prices to be \( p_t^* \). In any period, firms \( i \in [0, 1] \) have price \( p^*_t \) and firm \( i \in (1, 2] \) have price \( p^*_{t-1} \). The aggregate price is

\[
p_t = \frac{1}{2} p^*_t + \frac{1}{2} p^*_{t-1}.
\]

(13)

2.2.2. Pricing and information structure of rationally inattentive firms

When firms are subject to the information process constraint, they are unable to obtain perfectly accurate information about \( p_t^* \). Following earlier work by Sims (2003) and Mackowiak and Wiederholt (2009), the information structure is modeled by assuming that firms have limited “attention”, which implies that they cannot perfectly attend to all available information and have to choose what to attend to. Each firm is endowed with an information processing capacity \( K \geq 0 \) towards processing information about macroeconomic aggregates. Each individual firm obtains its own knowledge about \( p_t^* \) after processing relevant information. Idiosyncratic learning error may occur, resulting in information dispersion.

Firms’ profit-maximizing objective can be approximated by minimizing the expected profit loss in the two periods due to “pricing errors”, which is a quadratic difference between firms’ actual price and the profit-maximizing price:

\[
\min -E[(p_{it} - p_t^*)^2 + (p_{it} - p_{t+1}^*)^2|I_i].
\]

(14)

\( I_i \) is the information set of firm \( i \) in period \( t \), and \( p_{it} = p_{it}^0 \) is firm \( i \)’s optimally chosen price for period \( t \) and \( t + 1 \). The quadratic objective implies the following pricing rule:

\[
p_{it} = E\left[\frac{1}{2} p_t^* + \frac{1}{2} p_{t+1}^*|I_i]\right],
\]

(15)

or simply

\[
p_{it} = E[p_t^*|I_i].
\]

(16)

Firms’ optimal pricing strategy is to choose their price to be their rational expectation of the optimal reset price under full information, \( p_t^* \), given their knowledge about \( p_t^* \), the accuracy of which is determined by their information processing constraint:

\[
H(p_t^*|I_{t-1}) - H(p_t^*|I_t) \leq K,
\]

(17)

where \( H(X|Y) \) denotes the conditional entropy, a measure of the uncertainty that is involved in a random variable \( X \) when \( Y \) is known. For Gaussian distributed variables \( X \) and \( Y \), \( H(X|Y) = \frac{1}{2} \log_2 \left[ 2\pi e \det \Omega_{XY} \right] \), where \( \Omega_{XY} \) is the conditional covariance matrix. In the information constraint, \( H(p_t^*|I_{t-1}) \) measures the uncertainty of \( p_t^* \), before new information available at \( t \) is processed, and \( H(p_t^*|I_t) \) measures the uncertainty afterwards. Information processing capacity \( K \) limits the amount of uncertainty that can be reduced by firms’ information processing channel. The larger the information processing capacity, the more uncertainty firms can eliminate, and hence the better information they have on \( p_t^* \).

Firms receive a noisy signal \( s_{it} \) about \( p_t^* \). The quality of \( s_{it} \) improves with a larger \( K \), and the optimal structure of \( s_{it} \) is determined based on the structure of the objective and the distribution of the tracking variable \( p_t^* \). Firms’ posterior information set includes prior information, signals, and the parameter vector, i.e. \( I_i = (I_{t-1}, s_{it}), I_{it} = \Theta = (\sigma, \psi, \theta, \sigma_t, \gamma_t, \pi, \beta) \).

2.3. Government

The government consists of a fiscal authority and a monetary authority.

---

6 See Appendix B for the derivation.

7 The information structure is chosen optimally so that \( s_{it} \) is most informative of \( p_t^* \). For example, Sims (2003) shows that for a Gaussian white noise tracking target and a linear-quadratic objective, the optimal signal structure takes the form of “true state plus white noise error” and the information processing constraint can be reduced to a signal extraction problem with the variance of the noise inversely determined by \( K \).

The fiscal authority faces a budget constraint

\[ M_t + T_t + B_t \leq M_{t-1} + R_{t-1} B_{t-1} + \tau_t \int_0^1 p_{it} Y_{it} \, di \]  

(18)
in its taxation and bond issues. For simplicity, I assume that the fiscal authority fixes nominal bonds at a constant level, \( B_t = B \geq 0 \), and the proportional sales tax has a non-stochastic steady state value \( \tau \) that corrects the average distortion arising from monopolistic competition.

The monetary authority supplies money to maximize social welfare. I define the social welfare in a period as the utility of the representative household. Following similar steps as in Woodford (2003), Appendix C shows that a second-order approximation of household’s utility function is

\[ \Omega = -\operatorname{Var}(y_t - y_t^N)^2 - \delta \operatorname{Var}(y_{it} - y_t), \]  

where \( \delta = (\omega + \theta - 1)/(\sigma + \omega) \). \( \operatorname{Var}(x_{it}) = E_i(x_{it}^2) - (E_i x_{it})^2 \) denotes cross-sectional variance, where \( E_i x_{it} = \int x_{it} \, di \) is the cross-sectional mean. Throughout the analysis, I define the “optimal” monetary policy as the one that minimizes (19). The welfare expression is intuitive. The first term measures the variability of output around the efficient level. The second term captures the cross-sectional variability of output across different firms. Cross-sectional variability of output leads to inefficiency because it creates variability in labor supply around the efficient level. We can rewrite Eq. (19) using the demand equation (4) and market equilibrium conditions. The cross-sectional variance of output can be written in terms of the cross-sectional price dispersion:

\[ \Omega = -\operatorname{Var}(y_t - y_t^N)^2 - \omega \operatorname{Var}(p_{it} - p_t), \]  

(20)where \( \omega = \theta^2 \delta \). The welfare depends on the variance of output gap and the cross-sectional variance of relative prices. In this model, variability in relative price arises from Taylor staggered pricing and idiosyncratic pricing errors.

The central bank maximizes the social welfare by adjusting nominal demand. The central bank is assumed to have committed to a money supply rule before the realization of shocks. The timing of events is specified in Fig. 1.

The central bank commits to a rule that is contingent on the realization of fundamental shocks \( u_t \) and \( y_t^N \). I consider a simple description of the money supply rule:\footnote{That is, the fiscal authority pursues a Ricardian fiscal policy.}

\[ m_t = au_t + by_t^N. \]  

(21)It is assumed that the central bank have a large enough capacity to process all currently available information about the macroeconomic aggregates.\footnote{This simple rule allows for an analytical solution. Section 4 considers a more general setting in which the money supply is a general linear function of the fundamental shock sequence \( u_t \) and \( y_t^N \).}

2.4. Policy contingent pricing and information learning

The central bank’s choice of policy response coefficients \( a \) and \( b \) affect firms’ benchmark reset price and information learning. The optimal reset price under full information can be partitioned into two independent components regarding markup shocks and natural rate shocks respectively:\footnote{See Appendix D.}

\[ p_t = p^{m}_t + p^{\gamma}_t. \]  

(22)Each component follows a stationary Gaussian AR(1) process

\[ p^{m}_t = \phi p^{m}_{t-1} + \gamma \left( \frac{1}{a} - 1 \right) u_t, \]  

(23)\[ p^{\gamma}_t = \phi p^{\gamma}_{t-1} + \gamma \left( b - 1 \right) y_t^N, \]  

(24)where

\[ \phi = \frac{1 - \sqrt{a}}{1 + \sqrt{a}} \in (0, 1), \quad \gamma = \frac{2a}{1 + \sqrt{a}^2} \in (0, 1). \]

To acquire information about \( p_t \), firms need to acquire information about the shocks \( u_t \) and \( y_t^N \). I assume that firms process information about the two types of shocks in independent channels,\footnote{The reason is that the central bank specializes in processing information about the aggregate economy, and it has an obvious incentive to devote much of their massive resources to learn the information well. On the other hand, firms may be more constrained by their scale, and distracted by other trade-offs in their attention allocation. As studied in Mackowiak and Wiederholt (2009), it is optimal for firms to devote a majority of their attention to idiosyncratic conditions and only limited resources to the task of processing aggregate information.} with capacity \( K_u \) and \( K_y \) respectively.
Firms solve the following information processing problem:

$$\min_{k_u,k_y} E[(p_t - p_t^e)^2]$$

(25)

subject to

$$p_t = E[p_t^i | I_t].$$

(26)

$$p_t^i = p_t^{i,0} + p_t^{i,\gamma}.$$  
(27)

$$p_t^{i,0} = \phi p_t^{i,0 - 1} + \gamma \left( a + \frac{1}{a} \right) u_t,$$

(28)

$$p_t^{i,\gamma} = \phi p_t^{i,\gamma - 1} + \gamma (b - 1) y_t^{i,\gamma},$$

(29)

$$K_u \geq H(p_t^{i,\gamma} | I_{t-1}) - H(p_t^{i,\gamma} | I_{t-1}, \xi_t^{\gamma}),$$

(30)

$$K_y \geq H(p_t^{i,\gamma} | I_{t-1}) - H(p_t^{i,\gamma} | I_{t-1}, \xi_t^y),$$

(31)

$$K \geq K_u + K_y,$$

(32)

$$I_d = \{ \Theta; s_u, s_{i-1}, s_{i-2}, \ldots \},$$

(33)

$$s_u = (s_u^{i,0}, s_u^{i,\gamma}),$$

(34)

$$u_t \sim N(0, \sigma_u^2), \ y_t^{i,\gamma} \sim N(0, \sigma_y^2).$$

(35)

The objective (25) is a minimization problem of firms’ expected profit loss due to pricing mistake by choosing optimal attention allocation.\(^{13}\) Eq. (26) is the firms’ optimal pricing strategy. Eqs. (27)–(29) specify the partition of \(p_t^i\) and the distribution of its components. Eqs. (30)–(32) are firms’ information processing constraints. Eq. (33) specifies the information set. Eq. (34) gives a partition of signals. Finally, (35) specifies the distribution of shocks.

The problem is solved by a two-step procedure. In the first step, the optimal signal structure and optimal price are solved for a given attention allocation; in the second step, the optimal attention allocation is solved by evaluating the objective, taking into account its influence on firms’ pricing behavior. Appendix E gives the detailed solution procedure.

For a given attention allocation \(\{K_u, K_y\}\), the optimal pricing behavior for the updating firms \(i \in [0, 1]\) in period \(t\) is characterized by

$$p_t^i = p_t^{i,0} + p_t^{i,\gamma},$$

(36)

where

$$p_t^{i,0} = \gamma \left( a + \frac{1}{a} \right) \left\{ \sum_{l=0}^{\infty} \left[ \phi^l \left( 1 - \left( \frac{1}{2K_u} \right)^{l+1} \right) \right] u_{t-l} + \frac{1 - 2 - 2K_y}{2K_u - \phi^2} \sum_{l=0}^{\infty} \left( \phi \left( \frac{1}{2K_u} \right)^l \right) s_t^{i,\gamma - l - 1} \right\}$$

(37)

$$p_t^{i,\gamma} = \gamma (b - 1) \left\{ \sum_{l=0}^{\infty} \left[ \phi^l \left( 1 - \left( \frac{1}{2K_y} \right)^{l+1} \right) \right] y_{t-l}^{i,\gamma} + \frac{1 - 2 - 2K_y}{2K_y - \phi^2} \sum_{l=0}^{\infty} \left( \phi \left( \frac{1}{2K_y} \right)^l \right) s_t^{i,\gamma - l - 1} \right\}.$$  

(38)

\(s_t^{i,0} \sim i.i.d. N(0, \sigma_u^2)\) and \(s_t^{i,\gamma} \sim i.i.d. N(0, \sigma_y^2)\) are tracking errors. Eqs. (36)–(38) have several interesting implications. First, firms’ response to shock depends on the attention allocated to the type of shock. The more attention devoted to a shock, the stronger the price response when a shock occurs. Second, price heterogeneity arises even within the group of firms that can update prices because of the idiosyncratic tracking errors among firms. The amount of price dispersion depends on firms’ information processing capability. The more attention allocated to track a shock, the smaller the difference in firms’ information set and the smaller the price difference across firms. Furthermore, firms’ price response to shocks depends on the policy response parameters \(a\) and \(b\). A policy response that can fully offset the shock impact on the benchmark price \(p_t^i\).

\(^{13}\) Note that this objective is equivalent to profit maximization.
that is, $a = -1/\alpha$ and $b = 1$, eliminate the need for private firms to respond to the shocks. However, any accommodation to shocks will lead to a response by the private sector.

As the optimal reset price components are Gaussian AR(1) distributed and the objective is quadratic, the optimal signal takes the following form:14

$$s^u_t = p_t^{u0} + \sqrt{\frac{2 - 2K_u}{(1 - 2 - 2K_u)(1 - \phi^2 2 - 2K_u)}} s^u_t,$$

$$s^v_t = p_t^{v0} + \sqrt{\frac{2 - 2K_v}{(1 - 2 - 2K_v)(1 - \phi^2 2 - 2K_v)}} s^v_t.$$  

(39)

(40)

Note that the accuracy of the signals improves when greater information capacity is allocated to track the information.15

The value of the objective function when a firm chooses the optimal price specified in (36)–(38) is

$$E[(p_t - p_t^*)^2] = \frac{1 - \phi^2}{2K_u} Var(p_t^{u0}) + \frac{1 - \phi^2}{2K_v} Var(p_t^{v0})$$

$$= \frac{1}{2K_u - a} \left[ \gamma \left( a + \frac{1}{\alpha} \right) \sigma_u^2 + \frac{1}{2K_v - b} \gamma (b - 1)^2 \sigma_y^2 \right]$$

(41)

(42)

Firms’ profit loss is the sum of mean square error of the two tracking problems. The size of the loss component decreases with the information flow allocated to that tracking problem. With an exogenous and finite information processing capacity, firms face a tradeoff in their attention allocation. Firms’ profit loss also depends on the policy response parameter $a, b$. The more the optimal reset price is allowed to fluctuate, the larger the profit loss. Moreover, the volatility of the components of optimal reset price, $\sigma_u$ and $\sigma_v$ also affects profit loss since a more volatile target is more difficult to track and leads to larger errors.

Generally, the tracking errors $s^u_t$ and $s^v_t$ are correlated across firms.17 I write the tracking error as the sum of a common component and an idiosyncratic component:

$$s^u_t = \eta^u_t + \eta^u_i,$$

$$s^v_t = \eta^v_t + \eta^v_i,$$

(43)

(44)

where $\eta^u_t$ and $\eta^v_t$ are tracking errors common among firms, and $\eta^u_i$ and $\eta^v_i$ are idiosyncratic errors that disappear in aggregation, i.e., $\int_1^{1/2} \eta^u_idt = 0$, $\int_1^{1/2} \eta^v_idt = 0$. I assume that the volatility of the common component $\eta^u_t$ accounts for a fraction $\lambda_u$ of the variance of $s^u_t$ and the volatility of $\eta^v_t$ accounts for a fraction $\lambda_v$ of the variance of $s^v_t$. $\lambda_u$ and $\lambda_v$ can be affected by the degree of firms’ reliance on the common information sources and the informativeness of the sources. I assume $\lambda_u$ and $\lambda_v$ are exogenous parameters.

Aggregation across firms according to (11)18 gives the general price level

$$p_t = p_t^u + p_t^v,$$

(45)

where

$$p_t^u = \frac{\gamma}{2} \left( \alpha + \frac{1}{\alpha} \right) \left\{ k_u \gamma t \sum_{l=0}^{\infty} \phi^l (1 - (1 - k_u)^l) + k_v \gamma t \sum_{l=0}^{\infty} \phi^l (1 - (1 - k_v)^l) \right\}$$

$$+ \frac{\gamma}{2} \left( \alpha + \frac{1}{\alpha} \right) \left\{ k_u (1 - k_u)^2 \sum_{l=0}^{\infty} \phi^l (1 - (1 - k_u)^l) + k_v (1 - k_v)^2 \sum_{l=0}^{\infty} \phi^l (1 - (1 - k_v)^l) \right\}$$

(46)

(47)

$$p_t^v = \frac{\gamma}{2} (b - 1) \left\{ k_v \gamma t \sum_{l=0}^{\infty} \phi^l (1 - (1 - k_v)^l) + k_v (1 - k_v)^2 \sum_{l=0}^{\infty} \phi^l (1 - (1 - k_v)^l) \right\}$$

(48)

14 See Appendix E.

15 Average size of tracking errors $\sqrt{(2 - 2K_u)/(1 - 2 - 2K_u)(1 - \phi^2 2 - 2K_u)}$ and $\sqrt{(2 - 2K_v)/(1 - 2 - 2K_v)(1 - \phi^2 2 - 2K_v)}$ decrease with an increase in $K_u$ and $K_v$ respectively.

16 That is, $a$ is far from $-1/\alpha$ and the $b$ is far from 1.

17 It is not realistic that tracking errors are purely independent across firms. Firms share a need to acquire some macroeconomic variables and may rely on some common information sources, such as central bank announcements, available information releases by some private firms, etc. Hence firms’ information acquisition accuracy may be affected by these common sources in a similar way. Any systematic bias will lead to a tracking error that cannot be fully eliminated by aggregation.

18 $p_t$ is the price chosen by firms that can update price at period $t$, so $p_t^u = p_t^u, p_t^v$ is the price of the firms that update their price at period $t-1$, so $p_t^u = p_{t-1}^u$. The aggregate price can also be decomposed into two components as $p_t^u, p_t^v$, so that each consists of a series of one type of shock and the tracking errors in learning that type of shock.
where \( k_u \equiv 1 - 2^{-2K_u} \in [0,1] \) and \( k_y \equiv 1 - 2^{-2K_y} \in [0,1] \). Taking into account firms’ pricing behavior given \( \{K_u, K_y\} \) and considering the constraints that \( K_u, K_y \in [0, K] \) lead to the optimal attention allocation:

\[
K_u^* = \begin{cases} 
K & \text{if } R > \frac{2^K - \phi^2 2^{-K}}{1 - \phi^2} \\
\frac{1}{2} \log_2 \left( \frac{\phi^2 + 2^K R}{1 + \phi^2 2^{-K} R} \right) & \text{if } R \in \left[ \frac{1 - \phi^2}{2^K - \phi^2 2^{-K}}, \frac{2^K - \phi^2 2^{-K}}{1 - \phi^2} \right] \\
0 & \text{if } R < \frac{1 - \phi^2}{2^K - \phi^2 2^{-K}} 
\end{cases}
\] (50)

\[
K_y^* = K - K_u^*,
\] (51)

where \( R \) is the policy contingent volatility ratio between the two shocks, defined as

\[
R = \frac{\left( a + \frac{1}{\alpha} \right) \sigma_u}{\left( b - 1 \right) \sigma_y}.
\] (52)

Note that there is no heterogeneity in the attention allocation as the distribution of shocks and the monetary policy rules are common for all firms. The information problem is solved by equating the marginal value of information flow concerning markup shocks and concerning natural rate shocks at the interior solution. The key insight for the optimal attention allocation (50)–(52) is that firms’ optimal information choice is determined by policy contingent volatility ratio \( R \), which measures the relative size of the variation in \( p_t^m \), driven by markup shocks, and variation in \( p_t^n \), driven by natural rate shocks, conditional on a given policy response. A larger \( R \) increases the marginal value of additional information flow to track \( p_t^m \), so the information flow to track \( p_t^n \) is smaller at the interior solution. In general, firms’ optimal attention is determined by both the exogenous distribution of shocks and the policy response to the shocks. The larger the exogenous volatility is for the markup shocks relative to natural rate shocks, the more information flow is allocated to the conditions driven by markup shocks. Meanwhile, the optimal information choice also depends on the central bank’s policy choice. A more stabilized component of the optimal reset price requires less private attention concerning that component. To see this, an incentive for price stabilization to natural rate shocks, i.e. \( b \to 1 \), mitigates the variation in \( p_t^n \), and firms optimally switch their attention away from tracking the conditions driven by markup shocks.

On the other hand, a commitment of offsetting markup shocks, i.e. \( a \to -1/\alpha \), mitigates the variation in \( p_t^m \), and firms optimally switch their attention away from tracking the conditions driven by markup shocks.

3. A Ramsey problem

The central bank solves a Ramsey problem to maximize social welfare by choosing the policy response coefficients \( a \) and \( b \), taking the equilibrium conditions and firms’ optimal behavior as given.

\[
\max_{a,b} \Omega = -\text{Var}(y_t - y_t^{N})^2 - \omega \text{E}[\text{Var}(p_t - p_t)].
\] (53)

The set of constraints are (1) money supply rule (21); (2) cash-in-advance constraint \( y_t = m_t - p_t \); (3) full information benchmark reset price and its decomposition (27)–(29); firms’ optimal price and its decomposition (37) and (38); (4) aggregate price and its decomposition (45)–(49); (5) firms’ optimal allocation of information flow (50) and (51); and (6) Gaussian white noise distributed shocks (35).

To determine the optimal policy response, I need to consider how the terms in the welfare criterion depend on firms’ choices and policy response parameters \( \{a, b\} \).

3.1. Cross-sectional price dispersion

The cross-sectional variance of price dispersion is weighted in the welfare criterion. Using the results regarding firms’ prices and the aggregate price, I rewrite the cross-sectional variance of price dispersion as

\[
\text{E}[\text{Var}(p_t - p_t)] = \int_0^{1/2} (p_t^u - p_t^l)^2 \, dt + \int_{1/2}^1 (p_t^u - p_t^l)^2 \, dt
\]

\[
= \int_0^{1/2} (p_t^u - p_t^l)^2 \, dt + \int_{1/2}^1 (p_t^u - p_t^l)^2 \, dt + \int_0^{1/2} (p_t^l - p_t^l)^2 \, dt + \int_{1/2}^1 (p_t^l - p_t^l)^2 \, dt
\] (54)

\[
= \int_0^{1/2} (p_t^u - p_t^l)^2 \, dt + \int_{1/2}^1 (p_t^u - p_t^l)^2 \, dt + \int_0^{1/2} (p_t^l - p_t^l)^2 \, dt + \int_{1/2}^1 (p_t^l - p_t^l)^2 \, dt
\] (55)


\[ y^2 = \frac{1}{2} \left( a + 1 \right) \left\{ \frac{1}{1-\phi} \left( 1 - k_y^a \right)^2 + \frac{k_y^a(1-k_y^a)}{1-\phi^2(1-k_y^a)} \right\} \]  

\[ + \frac{1}{2} \sum_{i=0}^{\infty} \phi \left[ \left( 1 - (1-k_y)^{i+1} \right) \right] y_N^{i-1, -1} \]  

\[ f_y = m^t - p^t_y - y^N_t \]  

Proposition 1. In an economy with 2-period Taylor staggered pricing, optimal information processing by firms, and i.i.d. Gaussian distributed markup shocks and natural rate shocks, a money supply rule \( m_t = a u_t + b y_N^t \) that eliminates the cross-sectional variance of price dispersion satisfies:

\[ a = -\frac{1}{\alpha} \]  

\[ b = 1. \]

A policy that minimizes cross-sectional price dispersion should respond to natural rate shocks by the same magnitude, and offset markup shocks with a magnitude of \( 1/\alpha \) of the size of the shock. Note that the specified policy \( a = -1/\alpha \) and \( b=1 \) eliminates variation in the optimal reset price target \( p^r \) completely, i.e. \( \text{Var}(p^r) = 0 \). The intuition is as follows. Price dispersion in this economy is attributed to two sources. First, firms’ price contracts are staggered. Those that can update their prices in period \( t \) adjust their prices to be their optimal guess of the benchmark price \( p^t \), and those that cannot update their prices have prices that were their best guess of the benchmark price from last period, \( p^t_{t-1} \). Difference between \( p^t \) and \( p^t_{t-1} \) contributes to differences in the actual price chosen across firms. To minimize the price dispersion caused by staggered pricing, the central bank commits to a rule that stabilizes the path for the optimal reset price target in response to any contemporaneous shocks, i.e. \( p^t = p^t_{t-1} \), so firms, regardless of whether they can adjust their price or not, have the same tracking target in their information processing problem. Information dispersion is another source of price dispersion. Idiosyncratic errors arise from firms’ information processing, and firms’ knowledge of \( p^t \) and \( p^t_{t-1} \) may differ, which leads them to different prices. To minimize the information dispersion, the central bank should stabilize the optimal price target, as firms’ learning accuracy deteriorates with the variation of the target. Therefore, stabilizing optimal reset rate \( p^r \) serves both objectives well and eliminates both sources for cross-sectional price dispersion.

3.2. Variation in output gap

Following the independent assumption of the fundamental shocks, I decompose the variation of output gap based on the sources of fluctuation, and analyze the effect of policy responses on each component. That is,

\[ y_t - y^N_t = f_u(u_t, \{ \eta^N_t \}) + f_y(y^N_t) \]

where \( \{ f_u, f_y \} \) is a decomposition of the output gap. \( f_u \) summarizes the terms affected by markup shocks \( u_t \) and the learning errors arising from tracking markup shocks imperfectly, \( \{ \eta^N_t \} \), while \( f_y \) is affected by natural rate shocks \( \{ y^N_t \} \) and errors of tracking natural rate shocks \( \{ \eta^N_t \} \). I also rewrite the money supply rule to be \( m_t = m^t_u + m^t_y \), where \( m^t_u = a u_t \), and \( m^t_y = b y^N_t \).

The two components of the output gap takes the following functional form:

\[ f_y = m^t_y - p^y_t - y^N_t \]  

\[ = \left( 1 - \frac{1}{2} k_y \right) \left( b - 1 \right) y^N_t \]  

\[ - \frac{1}{2} \phi \left[ \sum_{i=0}^{\infty} \phi \left( 1 - (1-k_y)^{i+1} \right) y^N_i \right] \]  

\[ \text{intertemporal effect} \]

\[ \frac{1}{2} \phi \sqrt{\frac{1}{2} \phi \left( 1 - k_y \right) \sum_{i=0}^{\infty} \phi (1-k_y)^i \left( \eta^{N,i}_t + \eta^{N,i}_{t-1} \right) \]  

\[ \text{coordination effect} \]

\[ f_u = m^u_t - p^u_t \]

\[ \text{See Eq. (41). The mean square error increases with the variance of the tracking target.} \]
contemporaneous shock prevents past innovations from affecting current output. To see this, for any period response magnitude completely stabilizes the benchmark reset price path (natural rate shock component only) and on the current efficient output level, but they affect current optimal reset prices and actual output through sluggishly unchanged and aggregate output at its efficiency level.

output gap, the central bank should fully offset the contemporaneous effect of the natural rate shock, leaving current price variation. Intuitively, a natural rate shock moves the efficient output level by the same magnitude of the shock. To stabilize variation by a policy contingent magnitude of 

Proposition 2. In an economy with 2-period Taylor staggered pricing and optimal information processing by firms, a money supply rule $m_t = a\omega_t + by^N_t$ that minimizes the output gap variation caused by i.i.d. natural rate shock satisfies $b = 1$.

We can see why $b = 1$ minimizes the output gap variation caused by natural rate shocks by looking at each of the three components of $f_y$. First, contemporaneous natural rate shock $y^N_t$ can have immediate impact on the output gap, which calls for a change in money supply to mitigate the effect. From Eq. (62), a contemporaneous natural rate shock raises output gap variation by a policy contingent magnitude of $(b - 1)(1 - (1 - k_b) - (1 + \alpha))$. A policy rule that has $b = 1$ minimizes this variation. Intuitively, a natural rate shock moves the efficient output level by the same magnitude of the shock. To stabilize output gap, the central bank should fully offset the contemporaneous effect of the natural rate shock, leaving current price unchanged and aggregate output at its efficiency level.

Second, intertemporal shocks raise output variation due to price stickiness. Past natural rate innovations have no impact on the current efficient output level, but they affect current optimal reset prices and actual output through sluggishly evolving prices. To eliminate this inefficiency, the optimal policy should respond to the variation with a magnitude $b = 1$. This response magnitude completely stabilizes the benchmark reset price path (natural rate shock component only) and prevents past innovations from affecting current output. To see this, for any period $t$, the optimal reset price affected by natural rate shock takes a form
A policy rule of $b=1$ perfectly eliminates the effect of shocks $\{y_t^m\}$ on $\{p_t^m\}$, and it cuts off the intertemporal transmission of shock effects caused by staggered prices.

Third, minimizing the coordination effect requires the optimal policy response to natural rate shocks with a magnitude $b=1$. Based on the results on information processing, the accuracy of information that firms obtain improves when the variation of the tracking variable is smaller. A policy rule with $b=1$ offsets any shock impact on firms’ tracking target $\{p_t^m\}$. A constant target allows for perfect knowledge of the variable and eliminates tracking errors.

Combining all components, we see that $b=1$ achieves all three goals simultaneously without any policy tradeoff. Therefore, a complete price stabilization in response to natural rate shocks is optimal.

3.2.2. Output-stabilizing policy response to markup shocks

The policy analysis when markup shocks are present is more complicated. There are potential conflicting effects in the policy, and complete price stabilization is no longer optimal. Similarly to the analysis above, we can consider each of the three components of $f_u$ and the policy effect break-down specified in Eqs. (66)–(68).

Intertemporal markup shocks affect the output gap through staggered pricing. To mitigate the effect, the optimal policy response should maintain a stable price path, i.e. $a = -1/\alpha$, since the optimal reset price is

$$p^{m^0}_t = \phi p^{m^0}_{t-1} + a (a + 1) u_t. \quad (70)$$

A completely stabilized path of the optimal reset price eliminates the intertemporal transmission of shocks and prevents the output gap from being affected by past innovations in markups.

The coordination effect through firms’ imperfect information learning calls for a similar price stabilization objective. The size of tracking errors increases with the variation of optimal reset prices, so stabilizing $p^{m^0}_t$ facilitates information processing and minimizes the information dispersion and output gap variation. To stabilize $p^{m^0}_t$, the central bank should offset the markup shocks by setting $a = -1/\alpha$.

The standard effect, on the other hand, calls for a different objective. In response to contemporaneous markup shocks, an optimal policy must partially accommodate the shock and destabilize prices. A contemporaneous shock $u_t$ affects the variation of output gap by a policy contingent magnitude of $a - \gamma/(2a + 1/\alpha)k_u$, which calls for a positive response $a = \gamma k_u/((2 - \gamma k_u)\alpha)$ to stabilize the first component of $f_u$. Note that this magnitude increases in $k_u$. When firms are more attentive to markup disturbances, markup shocks lead to larger immediate adjustment by those that can reset their prices, which calls for a larger positive increase of money supply.

Therefore, in the presence of markup shocks, the central bank should consider the tradeoff between the policy effects. Complete stabilization is no longer optimal, and markup shocks lead to price dispersion and information dispersion. To stabilize the output gap, the monetary policy response to markup shocks should be a weighted average between the two suggested values by the objectives, as summarized by the proposition below.

**Proposition 3.** In an economy with 2-period Taylor staggered price setting and optimal information processing by firms, a money supply rule $m_t = a u_t + b y_t$ that minimizes the output gap variation caused by i.i.d. markup shocks is a weighted average of $-1/\alpha$ and $\gamma k_u/((2 - \gamma k_u)\alpha)$; that is, $a = (1 - \chi)\gamma k_u/((2 - \gamma k_u)\alpha + \chi(-1/\alpha))$, where $\chi \in (0, 1)$.

$\chi$ weights the importance of price stabilization in reducing the output gap variation. The analytical form of the weighting parameter $\chi$ is given in Appendix F. Fig. 2 explores the properties of $\chi$ with respect to the degree of price complementarity and information capacity.

![Fig. 2. Weighting the objectives in output gap stabilization.](image)
The key insight provide by Proposition 3 is that even if the welfare objective does not weight price stabilization directly, the output stabilization objective calls for price stabilization indirectly with a positive weight \( χ \), when firms are subject to price adjustment frictions and information frictions. Because of nominal frictions, some firms are unable to adjust their prices in response to markup variation and inefficiency arises, which calls for stabilizing the reset price path in response to markup shocks. Because of the information frictions, firms have an imperfect and diverse knowledge of market conditions, which leads to suboptimal pricing and production decisions. A commitment to stabilize prices facilitates information processing, reduces firms’ learning mistakes and aligns their expectations with the policy. According to Fig. 2, larger information inflow allocated to markup disturbances and lower strategic complementarities require greater emphasis of price stabilization by the central bank.

3.3. Optimal policy response

Based on Propositions 1 and 2, to minimize the cross-sectional price dispersion and output gap variation arising from natural rate shocks, the best policy response is to change the money supply one-for-one with the shock. Therefore, we can describe the optimal policy rule as follows:

\[
m_t = \alpha u_t + y^N_t.
\]

(71)

Optimal policy response to the markup shock \( \alpha^* \) is determined by weighting the minimizing price dispersion objective and output gap stabilization objective. Based on Propositions 1 and 3, I can derive an analytical solution for \( \alpha^* \).

Fig. 3 characterizes the main properties of \( \alpha^* \). It shows how the optimal magnitude of response to a one-unit increase in the markup shock varies with information capacity, with the rest of parameters taking standard values as in the literature: \( \sigma = \psi = 1 \), \( \theta = 2 \), which implies \( \alpha = 2/3 \) and the weighting parameter \( \alpha \) = 3. I also show how changes in the value for \( \alpha \) affect the solution. It is assumed that 50 percent of the variation of the tracking errors are common across firms, i.e. \( \tilde{h}_u = 1/2 \).

The determination of optimal policy response \( \alpha^* \) reflects the tradeoff between accommodating the standard effect of the markup shock and price stabilization. Note that a larger \( \alpha^* \) reflects more emphasis on accommodating the real effect of the markup shock and less emphasis on price stabilization. As shown in Fig. 3, the optimal response magnitude decreases with firms’ information capacity \( K \), suggesting greater emphasis on price stabilization. Intuitively, greater information flow concerning markup shocks leads to a stronger price response to the shocks by the firms that can reset their prices. It is thus more important to stabilize the optimal reset price path and reduce variation in firms’ prices. This emphasis outweighs the influence of an increase in \( K \) on other concerns, including an upward revision of the policy response to accommodate the standard effect of shock and smaller coordination effect.

The optimal response also depends on parameter \( \alpha \), the measure for the degree of strategic complementarity. A higher value of \( \alpha \) indicates weaker complementarities and a less persistent effect of markup shock on the optimal reset prices \( (\phi, \rho_u) \) decreases with \( \alpha \). Moreover, a larger \( \alpha \) also lowers the price stabilization benchmark \( \alpha = -1/\alpha \). The optimal response decreases with \( \alpha \) overall, especially when \( \alpha \) is near zero.

Fig. 4 compares the optimal policy response between staggered price and flexible price setting. I consider a flexible price model similar to Adam (2009). In the flexible price setting, firms can adjust price in every period, while other aspects of the model stay the same. The policy implications under the flexible price setting differ from the baseline model in mainly two aspects. First, flexible pricing eliminates the intertemporal effects of monetary policy and any persistent effect of i.i.d. shock; second, it allows for a stronger price response to markup shocks at the aggregate level, because of an increase in the response necessary to accommodate the standard effect of the shock. According to Fig. 4, staggered price setting induces a policy with more aggressive price stabilization when firms’ information processing capacity is low.

4. Optimal monetary policy response in a more general setting

This section extends the basic setting considered so far to a more general case. Shocks can be autocorrelated, and the monetary policy rule is a general linear function of the fundamental shocks. This setting allows for the consideration of policy response to past shocks, especially when shocks may have persistent effect on output gap variation and price dispersion. The main qualitative result is similar to the baseline model discussed in Section 3: the optimal response to a natural rate shock induces complete price stabilization, while for markup shocks, complete price stabilization is not optimal.

The shocks now evolves according to

\[
y^N_t = \rho_y y^N_{t-1} + \epsilon^y_t,
\]

\[
u_t = \rho_u u_{t-1} + \epsilon^u_t,
\]

(72)

(73)

where \( \epsilon^y_t \sim i.i.d. N(0, \sigma^2_y) \), \( \epsilon^u_t \sim i.i.d. N(0, \sigma^2_u) \), and \( \rho_u, \rho_y \in [0, 1] \).

---

22 See Appendix F.

23 The standard effect calls for a policy response \( a = \gamma k_u/(2 - \gamma k_u)\alpha \) increasing with \( k_u \).

24 The AR(1) distributed markup shocks and natural rate shock shocks can be obtained by assuming AR(1) distributed log productivity \( \{\ln A_t\} \) and log proportional tax \( \{\ln (1 - \tau_t)\} \).
4.1. A general policy rule

The central bank has the same welfare criterion
\[
\Omega = - \text{Var}(\gamma_t - y_t^N)^2 - \omega E[\text{Var}(\pi_t - \pi_t^0)].
\]  
(74)

It is assumed that the central bank now commits to a money supply rule that is a linear function of the fundamental shocks
\[
m_t = A(L)u_t + B(L)y_t^N,
\]  
(75)

where \( A(L) = a_0 + a_1 L + a_2 L^2 + \cdots \) and \( B(L) = b_0 + b_1 L + b_2 L^2 + \cdots \) are infinite-order lag polynomials.

The monetary policy rule of \( m_t = A(L)u_t + B(L)y_t^N \) and the innovation process (72) and (73) implies the following function form for the optimal reset price under full information:
\[
p_t^* = \frac{F(L)u_t + G(L)y_t^N}{p_t^*},
\]  
(76)

where the lag polynomial \( F(L) = f_0 + f_1 L + f_2 L^2 + \cdots \) and \( G(L) = g_0 + g_1 L + g_2 L^2 + \cdots \) are given by
\[
f_0 = \phi(f_0 + f_1) + \left(1 - \frac{1}{\alpha}\right) \left(a_0 + \frac{1}{\alpha}(1 + \rho_u) + a_1\right)
\]  
(77)

\[
f_j = \phi(f_{j-1} + f_{j+1}) + \left(1 - \frac{1}{\alpha}\right) (a_j + a_{j+1}), \quad \text{for } j = 1, 2, 3, \ldots
\]  
(78)
\[ g_0 = \phi(g_0 \rho_y + g_1) + \left( \frac{1}{2} - \phi \right) \left( (b_0 - 1)(1 + \rho_y) + b_1 \right) \]  
(79)

\[ g_j = \phi(g_{j-1} + g_{j+1}) + \left( \frac{1}{2} - \phi \right) (b_j + b_{j+1}), \quad \text{for } j = 1, 2, 3, \ldots \]  
(80)

where \( \phi = (1 - \alpha)/2(1 + \alpha) \). The components of the optimal reset price, \( p_t^* \) and \( p_t^{w*} \) are linear combinations of the markup shocks \( \{u_t\} \) and the natural rate shocks \( \{y_t^N\} \) respectively, and that the function form of \( p_t^* \) is contingent on monetary policy. Given that firms are tracking the optimal reset price and base their decisions on what they know about \( p_t^* \), the central bank's choice influences firms' behavior. In the meantime, the central bank chooses an optimal policy rule, taking as given its influence on the private sector.

Firms solve an information flow allocation problem

\[ \min_{k, x_y} E((p_t - p_t^*)^2) \]  
(81)

\[ p_t = E[p_t^* | l_t]. \]  
(82)

\[ p_t^* = F(L)u_t + G(L)y_t^N, \]  
(83)

\[ K_t^* \geq H(p_t^{w*} | l_{t-1}) - H(p_t^{w*} | l_t). \]  
(84)

\[ K_t^* \geq H(p_t^{w*} | l_{t-1}) - H(p_t^{w*} | l_t). \]  
(85)

\[ K^* \geq K_t^* + K_{t+1}. \]  
(86)

The central bank solves a similar Ramsey problem to maximize social welfare by choosing the policy response \( A(L) \) and \( B(L) \), taking the equilibrium conditions and firms' optimal behavior as given. The optimal policy is discussed in Sections 4.2 and 4.3.

4.2. Optimal response to natural rate shocks

The following proposition specifies the optimal policy response to natural rate shocks.

**Proposition 4.** In an economy with 2-period staggered pricing and optimal information processing by firms, a money supply rule \( m_t = A(L)u_t + B(L)y_t^N \), where \( A(L) = a_0 + a_1L + a_2L^2 + \cdots \) and \( B(L) = b_0 + b_1L + b_2L^2 + \cdots \) with the lag polynomial coefficients that satisfy

\[ b_0 + (1 + \rho_y) + b_1 = \phi(1 + \rho_y), \]  
(87)

\[ b_j + b_{j+1} = 0, \quad \text{for } j = 1, 2, 3, \ldots \]  
(88)

minimizes the output gap variation and cross-sectional price dispersion caused by natural rate shocks. At the optimal policy, \( p_t^{w*} = 0 \). In particular, a policy rule with

\[ b_j = \begin{cases} 1 & \text{for } j = 0 \\ 0 & \text{for } j = 1, 2, 3, \ldots \end{cases} \]  
(89)

satisfies the optimality condition.

The central bank can raise the money supply 1 to 1 in response to an increase in \( y_t^N \), which makes \( p_t^{w*} \) a constant. In firms' information problem, the optimal allocation of information flow depend on the relative volatility between \( p_t^* \) and \( p_t^{w*} \). When \( p_t^{w*} \) is perfectly stable, it is optimal for firms to ignore it and focus on \( p_t^* \) instead, i.e. \( K_t^* = 0 \). With zero information flow devoted to \( p_t^{w*} \), firms receive no information regarding natural rate shocks and their prices do not respond to natural rate shocks, leading to a perfectly stable path for the equilibrium aggregate price in response to natural rate shocks. Welfare is maximized in this situation, as there is no cross-sectional price dispersion, and aggregate output moves 1-1 with the efficient level of output. There is no tradeoff between stabilizing output and prices.

4.3. Response to markup shocks

I solve the optimal policy response to markup shock numerically. The reason is that the benchmark reset price generally does not follow an AR(1) process when the central bank follows a general money supply rule and the shocks are

\[ \text{See Appendix G for the derivation.} \]
autocorrelated. As firms’ reactions are policy contingent and the central bank considers firms’ reactions when chooses the money supply rule, solving the optimal policy response is a fixed point problem.

I compute the optimal policy response as follows. First, an initial guess concerning the policy rule is made. I then solve firms’ optimal pricing problem, taking as given the guessed policy rule and optimal response to natural rate shocks so that \( K^* = K \). Firms’ optimal price is obtained by taking the first order condition of firms’ quadratic minimization problem with the information processing constraint. I then compute the aggregate price and output. Second, I solve the central bank’s Ramsey problem and compute a money supply rule that maximizes the welfare, based on the results from the previous step. Third, I compare the money supply rule with the guess and update the guess until a fixed point is reached.

4.3.1. The benchmark case

The estimated process is \( m_t^e = \sum_{i=0}^{\infty} a_i u_{t-i} \). Since \( \{u_t\} \) follow a stationary AR(1) process, their effects decrease geometrically over time and eventually die out. The optimal monetary response to \( \{u_t\} \) should also be stationary and die out after a large number of periods. To solve the problem, I turn the infinite-dimensional problem into a finite-dimensional problem by restricting the number of lags in the moving average representation to 20, i.e. \( m_t^e = \sum_{i=0}^{20} a_i u_{t-i} \). I find that 20 lags can adequately approximates the policy response as adding additional lags has minimal effect.

Fig. 5 shows the optimal monetary policy response to a one-unit increase in the markup shock, with the standard parameters values \( \alpha = \frac{2}{3} \) and welfare weighting parameter \( \omega = 3.26 \).27 I pick the standard deviation of markup shocks \( \sigma_u \) to be 0.24, following Smets and Wouters (2007), and persistence \( \rho_u \) to be 0.89. The choice of shock volatility does not affect the solution since other second moments in the model can be normalized in terms of shock volatility. I will discuss how the changes in shock persistence \( \rho_u \) affect the result. I take that the total information capacity \( K = 0.5 \) to be the benchmark value, and Section 4.3.2 consider different levels of information capacity. One period is a quarter.28

26 \( \alpha = \frac{2}{3} \) and \( \omega = 3 \) follow \( \sigma = \psi = 1 \) and \( \theta = 2 \).

27 Woodford (2003) recommends a different range for \( \alpha \) between 0.1 and 0.15. Section 4.3.2 shows how changes in the value for \( \alpha \) affect the solution.

28 The two-period price duration is thus assumed to be 6 months. One could set the time period to be between 3.5 and 5.5 months to deliver the documented average price duration found in Bils and Klenow (2004) and Nakamura and Steinsson (2008).
The intuition is similar to Section 3. There are inefficiencies that monetary policy must offset in the presence of markup shocks. First, a positive markup raises the optimal reset price. Some firms are able to adjust their price to catch up with the change, but others cannot, which leads to greater price dispersion. Second, markup shocks lead to fluctuation in optimal reset prices. The information that firms obtain deteriorates with the volatility of optimal reset price. An increase in optimal reset price volatility increases information and price dispersion. When firms share some common information sources, their learning mistakes can cause the aggregate output to deviate from its efficient level. To reduce these two inefficiencies, the central bank can counteract the effect of the markup shocks and stabilize prices. However, reducing the real effect of shocks requires another consideration. Lowering the money supply to stabilize the optimal reset price amplifies the fall in output in the short run. Tradeoffs in the presence of markup shocks implies that complete price stabilization is not optimal.

At the optimal policy, output and prices are moving in the opposite directions, reflecting the policy tradeoff. The central bank must balance the need to accommodate the standard effect and to offset the intertemporal effect and coordination effect of the shocks. Fig. 5 shows that with standard model parameters, the optimal policy response leads to a sharp drop in money supply and output initially after the shock, reflecting the price smoothing emphasis. The reason is that markup shocks may cause particularly large variations in price dispersion and learning errors initially, so the need to offset the markup shocks and stabilize prices outweighs the need to stabilize output. It is therefore optimal for the central bank to emphasize price stabilization during this time. In the long run, when firms’ adapt to the innovation, it is optimal for the central bank to deemphasize price stabilization and focus more on accommodating the real effect of shocks. In Fig. 5, we see the optimal response increases in the long run and eventually dies out as the persistent effect of the shock dies out.

4.3.2. Effect of parameter values

I consider the effect of different parameter values on the optimal policy implication.

Fig. 6 shows how optimal policy is affected by firms’ information flow allocated to markup shocks. When firms can acquire more accurate information about markup shocks \((K = 1.5)\), the optimal policy response is stronger and less persistent than when firms have only very vague idea about the shocks \((K = 0.1)\). The policy implication is that the central bank needs to adjust the money supply more aggressively when firms are more attentive to markup shocks. Intuitively, when information capacity increases, firms are able to acquire better information about the shocks. When shocks occur, firms that are able to adjust their prices respond strongly to their perceived knowledge about the shocks. As a result, unless the central bank adopts a more aggressive monetary policy rule, a shock leads to a rise in cross-sectional price dispersion, since price contracts are staggered and tracking errors rise, and an increase in output variation since monetary policy has weaker real effects. Optimal policy thus emphasizes more on price stabilization initially. As illustrated in Fig. 6, when
attention is high ($K = 1.5$), the money supply falls more sharply initially, reflecting greater emphasis on price stabilization immediately after the shock, and quickly rebounds to a positive value to accommodate the real effect of shocks afterwards.

Figs. 7 and 8 illustrate the effect of strategic complementarities and shock persistence on optimal policy. We observe that the optimal response to a markup shock shifts more of its initial emphasis towards price stabilization when there is stronger

---

**Fig. 7.** Optimal policy response, markup shock persistence.

**Fig. 8.** Optimal policy response, strategic complementarities, persistent markup shocks.
complementarities (smaller $\alpha$) and more persistent shocks (larger $\rho_u$). Stronger complementarities and more persistent shocks lead to a more persistent path for the optimal reset price in response to markup shocks. Firms are able to track the optimal reset price better over time, which leads to stronger response by the updating firms. A markup shock thus can cause larger price dispersion and more variation in output gap, which induces a larger initial emphasis on price stabilization and more aggressive adjustment by the central bank.

5. Conclusion

This paper examines the policy implication for a rationally inattentive economy with nominal frictions in updating prices. The optimal monetary policy for this economy should fully offset the expected impact of natural rate shocks and completely stabilize prices, but complete price stabilization is not optimal in the presence of markup shocks. The central bank should carefully balance the need to accommodate the real effects of markup shocks and to offset the shock in order to maintain a stable price path, taking into account the coordination effect and the reduction of cross-sectional price dispersion due to information dispersion and price rigidity. In the short-run, the emphasis on price stabilization outweighs the emphasis on output stabilization after the shock, as stabilizing prices initially improves firms’ knowledge and reduces the inefficiency caused by price stickiness and information dispersion.

Acknowledgment

I thank Paul Evans, Bill Dupor, Pok-sang Lam and Aaron Popp for valuable advice and support in this project. I also thank the editor and two anonymous referees for important comments and suggestions. Additional thanks are due to the seminar participants at The Ohio State University.

Appendix A. Profit-maximizing price under perfect information

A.1. Deriving the profit-maximizing price under perfect information

The representative household’s utility maximization problem with respect to consumption of good $C_i$ leads to the demand for good $i$ in household $i$’s consumption bundle $C_i = (P_i/P)^{-\theta}C$, and the optimal consumption-labor decision leads to $W_t/P_t = C_i^L$. Market clearing condition satisfies $C_i = Y_t$. $C = Y$.

When firms have perfect information, their profit maximization problem is specified as follows:

$$\text{max}_{\tilde{P}_t} \tilde{H}_t = (1 - \tau_t) \left( \frac{P_t}{\tilde{P}_t} \right) Y_t - \left( \frac{W_t}{\tilde{P}_t} \right) L_t, \quad (A.1)$$

s.t. $L_t = \frac{Y_t}{A_t}, \quad (A.2)$

$$Y_t = \left( \frac{P_t}{\tilde{P}_t} \right)^{-\theta} Y_t, \quad (A.3)$$

$$W_t/P_t = Y_t^\theta L_t^\psi. \quad (A.4)$$

An individual firm takes $Y_t, A_t, \tau_t, P_t, W_t$ as exogenous.

The optimal price $P^*_t$ satisfies

$$\frac{P^*_t}{P_t} = \frac{1}{\theta - (1 + \psi)A_t^\psi} \exp\left( (-1 + \psi)Y_t^\psi \right). \quad (A.5)$$

After taking logs, I obtain

$$p^*_t = p_t + \alpha_t - \frac{1 + \psi}{1 + \theta\psi} \ln(1 - \tau_t) + \frac{1}{1 + \theta\psi} \ln\left( \frac{\theta}{\theta - 1} \right). \quad (A.6)$$

where $\alpha = \sigma + \psi / (1 + \theta\psi)$, $\alpha_t = \ln(A_t)$. Note that the profit-maximizing price is the same across households. Denote this common profit-maximizing price as $p^*$. The natural rate of output is defined as the output level when $p_t = p^*, \forall i$, and the sales tax rate is at its average level $\tau$ (no markup shocks). Substituting the definition of shocks

$$Y^N_t = \frac{1}{\psi + \sigma} \left( (1 + \psi) \alpha_t - \ln\left( \frac{\theta}{\theta - 1} \right) - \ln(1 - \tau) \right)$$

and

$$u_t = \frac{1}{1 + \theta\psi} \ln\left( \frac{1 - \tau}{1 - \tau_t} \right).$$
in the above equation yields

\[ p^*_t = p_t + \alpha(y_t - y^*_t) + u_t. \] (A.7)

### Appendix B. Firms' quadratic objective

This appendix shows that firms' profit-maximizing objective can be approximated by minimizing a quadratic form of pricing errors.

The real profit function can be rewritten in terms of the innovation variables. The quadratic approximation of the real profit function around the steady state \((0,0,0,0,0)\) takes the form

\[
\hat{\Pi}(p_t, w_t, y_t, a_t, \tau_t) = \hat{\Pi}(0,0,0,0,0) + \hat{\Pi}_1 p_t + \hat{\Pi}_2 w_t + \hat{\Pi}_3 y_t + \hat{\Pi}_4 a_t + \hat{\Pi}_5 \tau_t
\]

\[ + \frac{1}{2} \hat{\Pi}_{11} p^2_t + \hat{\Pi}_{12} p_t w_t + \hat{\Pi}_{13} p_t y_t + \hat{\Pi}_{14} p_t a_t + \hat{\Pi}_{15} p_t \tau_t \]

\[ + \text{other cross-product terms without } p_t. \] (B.3)

where \(\tau_t \equiv \tau_t - \tau_t \{\hat{\Pi}_{11}, \ldots \hat{\Pi}_{15}\}\) and \(\{\hat{\Pi}_{11}, \ldots \hat{\Pi}_{15}\}\) are first and second order of derivatives of the function \(\hat{\Pi}(p_t, w_t, y_t, a_t, \tau_t)\) with respect to their corresponding argument, evaluated at the steady state.

Let \(\hat{\Pi}(p^*_t, w_t, y_t, a_t, \tau_t)\) be the benchmark high profit level, which is achieved when firm correctly sets its price to be the full information benchmark profit-maximizing price \(p^*_t\). Let \(p_t\) be any other price chosen by the firm.

\[
\hat{\Pi}(p^*_t, w_t, y_t, a_t, \tau_t) - \hat{\Pi}(p_t, w_t, y_t, a_t, \tau_t) = \hat{\Pi}_{11} (p^*_t - p_t)^2 + \hat{\Pi}_{12} (p^*_t - p_t) w_t + \hat{\Pi}_{13} (p^*_t - p_t) y_t + \hat{\Pi}_{14} (p^*_t - p_t) a_t + \hat{\Pi}_{15} (p^*_t - p_t) \tau_t
\]

\[ + \frac{1}{2} \hat{\Pi}_{11} (p^*_t - p_t)^2. \] (B.4)

Steady state variable values satisfy \(\hat{\Pi}_1 = 0\). The first order condition with respect to the first argument implies

\[ \hat{\Pi}_{11} p^*_t + \hat{\Pi}_{12} w_t + \hat{\Pi}_{13} y_t + \hat{\Pi}_{14} a_t + \hat{\Pi}_{15} \tau_t = 0. \]

Substituting \(\hat{\Pi}_1 = 0\) and \(\hat{\Pi}_{11} p^*_t + \hat{\Pi}_{12} w_t + \hat{\Pi}_{13} y_t + \hat{\Pi}_{14} a_t + \hat{\Pi}_{15} \tau_t = 0\) into the above equation, the profit difference is quadratic in the price difference:

\[
\hat{\Pi}(p^*_t, w_t, y_t, a_t, \tau_t) - \hat{\Pi}(p_t, w_t, y_t, a_t, \tau_t) = -\frac{1}{2} \hat{\Pi}_{11} (p^*_t - p_t)^2 - \frac{1}{2} \hat{\Pi}_{11} p^*_t + \hat{\Pi}_{11} p^*_t p_t
\]

\[ = -\frac{1}{2} \hat{\Pi}_{11} (p^*_t - p_t)^2. \] (B.5)

where \(\hat{\Pi}_{11} < 0\).

Therefore, minimizing a quadratic form of price deviation from the profit-maximizing price is equivalent to maximizing real profit.

The model discussed in this paper assumes Taylor staggered price setting, so firms need to choose a price \(p_t\) for both periods. It is assumed that firms value next period equally as today. The profit-maximization problem of firms can be approximated by a two-period quadratic minimization problem of pricing errors:

\[
\min -E[(p_t - p^*_t)^2 + (p_{t+1} - p^*_{t+1})^2]. \] (B.8)

\(p^*_t\) is the profit-maximizing price for period \(t\), and \(p^*_{t+1,1} \equiv E(p^*_{t+1} | I_t)\) is the rational expectation of period \(t+1\)'s profit-maximizing price \(p^*_{t+1}\) using all the information available at period \(t\), both exogenous to individual firms. Subtracting an exogenous term \(\frac{1}{2} (p^*_t - p^*_{t+1})^2\) from Eq. (B.8) and scale it by 1/2 yields

\[
\min -E[(p_t - p^*_t)^2]. \] (B.9)

where \(p^*_t = \frac{1}{2} p_t + \frac{1}{2} p^*_{t+1}.\)

### Appendix C. Welfare criterion

This part of derivation is similar to Ball et al. (2005). I approximate the utility function around the steady state when real disturbances are at their means. As Woodford (2003) emphasizes, it is important for the accuracy of the approximation that the steady state is close to being efficient. I assume that \(E(\theta / \theta - 1 + (1 - \tau)) = 1\), which implies

\[
y_t = \frac{1 + \psi}{\sigma + \psi} \tau_t, \]

where \(\tau_t = \log (\tilde{A}_t).\)

The welfare function

\[
U_t = \frac{C_t^{1-\sigma} - 1}{1 - \sigma} - \frac{L_t^{1+\psi}}{1 + \psi} \]

\[+ \text{other terms}. \] (C.2)
\[ Y_t^{1-\sigma} - 1 = \frac{1}{1+\psi} \left( \int_{0}^{1} \left( \frac{Y_t}{A_t} \right) \right)^{1+\psi} \]

\[ = e^{1-\sigma\gamma_t} \left( 1 + \psi \left( \int_{0}^{1} e^{\gamma_{at} - \epsilon_t} \right) \right)^{1+\psi} - 1 + \frac{1}{1+\psi} \left( \int_{0}^{1} \left( e^{\gamma_{at} - \epsilon_t} \right) \right)^{1+\psi} \]

Define \( \hat{Y}_t = Y_t - \bar{\gamma}_t \), \( \check{Y}_t = y_t - \bar{\gamma}_t \), and \( \hat{a}_t = a_t - \bar{\gamma}_t \). A second order Taylor approximation of the aggregate utility around the steady state \( (\hat{Y}_t, \hat{a}_t, \bar{\gamma}_t) \) gives

\[ \Omega_t = e^{1-\sigma\gamma_t} \left( \hat{Y}_t + \frac{1-\sigma}{2}\hat{\gamma}_t^2 \right) - e^{1+\psi(\gamma_t - \bar{\gamma}_t)} \left( \frac{1+\psi}{2} \hat{Y}_t^2 - (1+\psi)\hat{Y}_t\hat{a}_t \right) \]

\[ \Omega_t = \left[ e^{1-\sigma\gamma_t} \left( \hat{Y}_t + \frac{1-\sigma}{2}\hat{\gamma}_t^2 \right) - e^{1+\psi(\gamma_t - \bar{\gamma}_t)} \left( \frac{1+\psi}{2} \hat{Y}_t^2 - (1+\psi)\hat{Y}_t\hat{a}_t \right) \right] + t.i.p. \]

where \( \Omega_t = U_t - U_t \). Additional terms that are independent of policy are collected in \( t.i.p. \)

Defining the cross-sectional mean as \( E_i(\hat{Y}_t) = \int \hat{Y}_t \, dt \) and the cross-sectional variance as \( \text{Var}_{i}(\hat{Y}_t) = E_i(\hat{Y}_t^2) - E_i(\hat{Y}_t)^2 \), we can rewrite this expression as

\[ \hat{y}_t = \hat{E}_i(\hat{Y}_t) - \frac{1}{2\theta} \hat{\text{Var}}_{i}(\hat{Y}_t). \]

Using the above equation to substitute for \( E_i(\hat{Y}_t) \) in the previous expression for \( U_t \), we obtain

\[ \Omega_t \approx e^{1-\sigma\gamma_t} \left( \hat{y}_t + \frac{1-\sigma}{2}\hat{\gamma}_t^2 \right) - e^{1+\psi(\gamma_t - \bar{\gamma}_t)} \left[ \hat{y}_t + \frac{1+\psi}{2} \hat{y}_t^2 + \frac{1+\psi\theta}{2\theta}\hat{\text{Var}}_{i}(\hat{Y}_t) - (1+\psi)\hat{a}_t\hat{y}_t \right]. \]

Note that \( (1-\sigma)\gamma_t = (1+\psi)(\gamma_t - \bar{\gamma}_t) \). Combining the two terms in brackets we obtain

\[ \Omega_t \approx - e^{1-\sigma\gamma_t} \frac{\sigma + \psi}{2} \left( \hat{y}_t - \hat{\gamma}_t\hat{y}_t + \frac{1+\psi}{\sigma + \psi} \hat{\text{Var}}_{i}(\hat{Y}_t) \right) \]

Next, recall the definition of the natural rate \( \hat{y}_N = (1+\psi)/(\sigma + \psi) \hat{\gamma}_t \) to substitute \( \hat{a}_t \), add a term \( (\hat{y}_N^2)^2 \) and drop the proportionality factor, which is beyond the control of the central bank, gives

\[ \Omega_t \approx - \hat{y}_t \hat{y}_N^2 - \hat{Y}_t \hat{y}_N^2 - E(\hat{y}_N^2 - \hat{y}_N) \hat{y}_N - \hat{y}_t \hat{y}_N \]

\[ \Omega_t \approx - \hat{y}_t \hat{y}_N^2 - E(\hat{y}_N^2 - \hat{y}_N) \hat{y}_N - \hat{Y}_t \hat{y}_N^2 - \hat{y}_t \hat{y}_N \]

\[ \Omega_t \approx - \hat{y}_t \hat{y}_N^2 - E(\hat{y}_N^2 - \hat{y}_N) \hat{y}_N - \hat{Y}_t \hat{y}_N^2 - \hat{y}_t \hat{y}_N \]

We can take the unconditional expectation of the above equation to obtain

\[ \Omega = \text{Var}(y_t - \hat{y}_N^2)^2 + \delta \text{Var}(y_t - \hat{y}_t) \]

where \( \delta = (\theta^{-1} + \psi)/(\sigma + \psi) \). Use the demand function \( y_t - y_t = -\theta(p_t - p_t) \).

\[ \Omega = \text{Var}(y_t - \hat{y}_N^2)^2 + \omega \text{Var}(p_t - p_t) \]

where \( \omega = \delta^2 \). The expectation of the sum of discount utility over time equals this expression.

**Appendix D. Optimal reset price under full information and uncorrelated shocks**

**Appendix A** has shown that the profit-maximizing price for period \( t \) under full information takes the following form:

\[ p_t^* = p_t + \alpha(y_t - \hat{y}_N^2) + u_t \]

Substituting \( y_t \), using the nominal demand \( m_t = p_t + y_t \) in the above expression yields

\[ p_t^* = (1-\alpha)p_t + am_t + u_t - \alpha\hat{y}_N^2 \]
The central bank supplies money according to a rule \( m_t = au_t + by_t^1 \), the profit-maximizing price can be rewritten as

\[
p_t^* = (1 - \alpha)p_t + a \left( \alpha + \frac{1}{\alpha} \right) u_t + a(b - 1)y_t^N. \tag{D.3}
\]

When firms have perfect information, the group of firms that can update their price chooses

\[
p_t^* = \frac{1}{2} p_t^* + \frac{1}{2} p_{t+1}^*, \tag{D.4}
\]

and the group that cannot update price in period \( t \) has a price optimally chosen in the last period, \( p_{t-1}^* = \frac{1}{2} p_{t-1}^* + \frac{1}{2} p_{t-2}^* \). Full information symmetric equilibrium requires the aggregate price to be the average between two pricing group’s optimal reset price:

\[
p_t = \frac{1}{2} p_t^* + \frac{1}{2} p_{t-1}^*. \tag{D.5}
\]

Eqs. \( \text{(D.4)} \) and \( \text{(D.3)} \) imply

\[
p_t^* = \frac{1}{2} p_t^* + \frac{1}{2} p_{t+1}^* \tag{D.6}
\]

\[
p_t^* = \frac{1}{2} (1 - \alpha)(p_t + p_{t+1}) + \frac{1}{2} \alpha \left( \alpha + \frac{1}{\alpha} \right) u_t + \frac{1}{2} a(b - 1)y_t^N. \tag{D.7}
\]

The second line uses the fact that the shocks are i.i.d. so that \( u_{t+1} = 0 \) and \( y_{t+1}^N = 0 \). Substituting the aggregate price using Eq. \( \text{(D.5)} \) and rearranging to have \( p_t^* \) on the left hand side yields

\[
p_t^* = D(p_{t-1}^* + p_{t+1}^*) + Z_t, \tag{D.8}
\]

where \( D \equiv \frac{1}{2} (1 - \alpha)/(1 + \alpha) \) and \( Z_t \equiv (\frac{1}{2} - D)[(a + 1/\alpha) u_t + (b - 1)y_t^N] \). Use the distribution of shocks, it follows that \( E_t Z_{t+1} = E_t Z_{t+2} = \cdots = 0 \). I can rewrite the above expression using lag operators. Here I use the rule that when the lag operator is applied to a term involving expectations, it lags the date of the variable, then \( p_{t+1}^* = L^{-1} p_t^* \). Eq. \( \text{(D.8)} \) can be rewritten as

\[
p_t^* = D(Lp_{t-1}^* + L^{-1} p_t^*) + Z_t, \tag{D.9}
\]

or

\[
(1 - DL - DL^{-1}) p_t^* = Z_t. \tag{D.10}
\]

Note that \( 1 - DL - DL^{-1} \) can be factored as \( (1 - \phi L^{-1})(1 - \phi L)_D \), where \( \phi = (1 - \sqrt{\alpha})/(1 + \sqrt{\alpha}) \). Thus we have

\[
(1 - \phi L^{-1})(1 - \phi L)_D = \frac{D}{\phi} Z_t. \tag{D.11}
\]

Multiplying both sides by \( (1 - \phi L^{-1})^{-1} = 1 + (L^{-1} + \phi L^{-2} + \phi^2 L^{-3} + \cdots) \) in the above equation yields

\[
(1 - \phi L)p_t^* = \frac{D}{\phi} \left( 1 + \phi L^{-1} + \phi^2 L^{-2} + \phi^3 L^{-3} + \cdots \right) Z_t. \tag{D.12}
\]

Note that \( L^{-1} Z_t = L^{-2} Z_t = L^{-3} Z_t = \cdots = 0 \). The expression can be simplified as

\[
(1 - \phi L)p_t^* = \frac{D}{\phi} Z_t, \tag{D.13}
\]

or

\[
p_t^* = \phi p_{t-1}^* + \frac{D}{\phi} Z_t \tag{D.14}
\]

\[
p_t^* = \phi p_{t-1}^* + \frac{D}{\phi} \left( \frac{1}{2} - D \right) \left[ (a + 1/\alpha) u_t + (b - 1)y_t^N \right]. \tag{D.15}
\]

Define \( \gamma \equiv D/\phi (\frac{1}{2} - D) = 2\alpha/(1 + \sqrt{\alpha})^2 \) and decompose \( p_t^* \) according to the source of fluctuation yields Eqs. \( \text{(28)} \) and \( \text{(29)} \).

**Appendix E. Solving firms’ information problem with Gaussian AR(1) target**

The following proposition specifies the signal structure and the solution to a quadratic tracking problem with a Gaussian AR(1) tracking target.

**Proposition 5.** For a stationary Gaussian AR(1) distributed tracking variable

\[
X_t = \phi X_{t-1} + e_t, \tag{E.1}
\]
with \( e \sim \text{i.i.d. } N(0, \sigma_e^2) \) and \( \phi \in [0, 1) \), an optimal information learning problem with quadratic objective \( E(\hat{X}_t - X_t)^2 \) and information capacity \( K \) implies the following the optimal signal structure and response:

\[
s_t = X_t + 2^{-K} \sqrt{\frac{1}{(1-2^{-2K})(1-\phi^2 2^{-2K}) \eta_t}}
\]  

(E.2)

\[
\hat{X}_t = \sum_{i=0}^{\infty} \phi^i \left( 1 - \left( \frac{1}{2K} \right)^{i+1} \right) e_{t-i} + \sqrt{\frac{1-2^{-2K}}{2^{-2K} - \phi^2}} \sum_{i=0}^{\infty} \left( \frac{\phi}{2^{-2K}} \right)^i \eta_{t-i}.
\]  

(E.3)

with stochastic tracking noise \( \eta \sim \text{i.i.d. } N(0, \sigma_e^2) \). The value of the objective function at optimal learning is

\[
E(\hat{X}_t - X_t)^2 = \frac{\sigma_e^2}{2^{-2K} - \phi^2}.
\]  

(E.4)

The proof is based on Sims (2003) and Mackowiak and Wiederholt (2009). The idea is to derive a lower bound for the quadratic tracking errors \( E(X_t - \hat{X}_t)^2 \), and then show the \( X_t \) and signal \( S_t \) attain this bound.

The information constraint can be written as

\[
\frac{1}{2} \log_2 \left( \frac{\text{Var}(X_t|\hat{X}_t^t)}{\text{Var}(X_t|\hat{X}_t^t)} \right) \leq K.
\]  

(E.5)

The AR(1) property of \( X_t \) implies

\[
\text{Var}(X_t|\hat{X}_t^{t-1}) = \phi^2 \text{Var}(X_{t-1}|\hat{X}_t^{t-1}) + \sigma_e^2
\]  

(E.6)

\[
\text{Var}(X_t|\hat{X}_t^{t-1}) = \phi^2 E(\hat{X}_{t-1} - X_{t-1})^2 + \sigma_e^2.
\]  

(E.7)

Stationarity requires

\[
E(\hat{X}_t - X_t)^2 = E(\hat{X}_{t-1} - X_{t-1})^2
\]  

(E.8)

Using this result in the information constraint, I obtain

\[
\frac{1}{2} \log_2 \left( \frac{\phi^2 E(\hat{X}_t - X_t)^2 + \sigma_e^2}{E(\hat{X}_t - X_t)^2} \right) \leq K,
\]  

(E.9)

which implies a lower bound

\[
E(\hat{X}_t - X_t)^2 \leq \frac{\sigma_e^2}{2^{-2K} - \phi^2}.
\]  

(E.10)

The specific \( \{ \hat{X} \} \) achieves this lower bound. To see this

\[
X_t - \hat{X}_t^t = \frac{\phi}{2^{-2K}} (X_{t-1} - \hat{X}_{t-1}^t) + \frac{1-2^{-2K}}{2^{-2K} - \phi^2} e_t - \sqrt{\frac{1-2^{-2K}}{2^{-2K} - \phi^2}} \eta_t.
\]  

(E.11)

which implies

\[
E(\hat{X}_t^t - X_t)^2 = \frac{\sigma_e^2}{2^{-2K} - \phi^2}.
\]  

(E.12)

Applying Proposition 5 to firms’ problem in learning optimal reset prices \( p_t^u \) and \( p_t^y \), we can obtain the solution for firms’ optimal prices \( p_t^u \) and \( p_t^y \) specified in Eqs. (37) and (38), signal structure (39) and (40), and the expected profit loss (41). Aggregating individual firms’ price, specified in Eqs. (37) and (38), across firms yields the aggregate price (47)–(49).

The optimality condition for information learning requires that the marginal benefit of learning the markup shocks equates to the marginal benefit of learning natural rate shocks for attention allocation \( (K_u, K_y) \) for an interior solution. That is,

\[
\frac{\partial E(p_t^u - p_t^\phi)^2}{\partial K_u} = \frac{\partial E(p_t^y - p_t^\phi)^2}{\partial K_y}.
\]  

(E.13)

Solving the above equation taking as given Eqs. (41), (47)–(49) and the constraint \( K_u + K_y \leq K \) yields the optimal attention allocation (50) and (51).

Appendix F. The weighting parameter for optimal response to uncorrelated markup shocks

The optimal response is a weighted average between three objectives.

---

29 Let \( \hat{X}_t \) be \( p_t^y \) and \( X_t \) be \( p_t^u \) in the tracking problem for \( p_t^y \), and \( \hat{X}_t \) be \( p_t^u \) and \( X_t \) be \( p_t^y \) in the tracking problem for \( p_t^u \).
The weight for the first objective of accommodating the effect of contemporaneous shock is

$$w_1 = \left(1 - \frac{\gamma k_u}{2}\right)^2.$$  \hfill (F.1)

The weight for the second objective of offsetting the effect of intertemporal shock is

$$w_2 = \frac{\gamma^2}{4} \left[1 + \phi - \frac{1}{1 - \phi} \frac{1 - k_u(1 + \phi(1 - k_u))}{1 - \phi(1 - k_u)}\right]^2.$$  \hfill (F.2)

The weight for the third objective of offsetting the effect caused by tracking errors is

$$w_3 = \frac{\gamma^2 \lambda^2}{2} \frac{k_u(1 - k_u)}{1 - (1 - k_u)\rho^2} \left[1 - \frac{1}{1 - \phi(1 - k_u)}\right]^2.$$  \hfill (F.3)

The weighting parameter $\chi = (w_2 + w_3)/(w_1 + w_2 + w_3)$.

The weight for minimizing cross-sectional price dispersion in the welfare criterion is

$$w_4 = \omega\rho^2 \left[\frac{1}{1 - \phi} \frac{1 - k_u}{1 - (1 - k_u)\rho^2} \left(1 - \frac{\lambda_u}{1 - \phi(1 - k_u)}\right)^2 + \frac{2k_u(1 - k_u)}{1 - (1 - k_u)\rho^2} \left(1 - \frac{\lambda_u}{1 - \phi(1 - k_u)}\right)^2 \right].$$  \hfill (F.4)

The analytical solution for optimal response coefficient to markup shocks is

$$a^* = \frac{w_1}{w_1 + w_2 + w_3 + w_4} \left(\frac{\gamma k_u}{2 - \gamma k_u\alpha}\right) - \left(1 - \frac{w_1}{w_1 + w_2 + w_3 + w_4}\right) \frac{1}{\alpha}.$$  \hfill (F.6)

### Appendix G. Optimal reset price under a general policy and correlated shocks

This part derives the function form for optimal reset price under full information $p^*_t$, when the money supply rule takes a moving average representation of shocks. The optimal reset price also takes a moving average representation:

$$p^*_t = F(L)u_t + G(L)y^N_t,$$  \hfill (G.1)

where the lag polynomials $F(L) = f_0 + f_1 L + f_2 L^2 + \ldots$ and $G(L) = g_0 + g_1 L + g_2 L^2 + \ldots$ are policy contingent.

Appendix A shows that the profit maximizing price takes the form of $p^*_t = p_t^* + \alpha(y_t - y^N_t) + u_t$. The cash in advance constraint $m_t = p_t^* + y_t$, and the profit-maximizing price equation imply

$$p^*_t = \frac{1}{2} p_t^* + \frac{1}{2} p^*_{t+1|t}.$$  \hfill (G.2)

$$p^*_t = \frac{1}{2} \left[1 - \alpha \right] \left[p_t + p_{t+1|t}\right] + \alpha \left(m_t + m_{t+1|t}\right) - \alpha \left(y_t + y_{t+1|t}\right) + \left(u_t + u_{t+1|t}\right).$$  \hfill (G.3)

The aggregate price in the full information symmetric equilibrium is

$$p_t = \frac{1}{2} p_t^* + \frac{1}{2} p^*_{t-1}.$$  \hfill (G.4)

The central bank sets the money supply in period $t$ according to the following rule in order to maximize social welfare:

$$m_t = A(L)u_t + B(L)y^N_t,$$  \hfill (G.5)

where $A(L) = a_0 + a_1 L + a_2 L^2 + \ldots$ and $B(L) = b_0 + b_1 L + b_2 L^2 + \ldots$ The stationary AR(1) processes for $\{u_t\}$ and $\{y^N_t\}$ implies the rational forecasting $y_{t+1|t} = \rho_y y_t$, and $u_{t+1|t} = \rho_a u_t$. The rational expectation for money supply in $t+1$ thus takes the following form:

$$m_{t+1|t} = \hat{A}(L)u_t + \hat{B}(L)y^N_t,$$  \hfill (G.6)

where $\hat{A}(L) = (a_0\rho_y + a_1) + a_2 L + a_3 L^2 + a_4 L^3 + \ldots$ and $\hat{B}(L) = (b_0\rho_y + b_1) + b_2 L + b_3 L^2 + b_4 L^3 + \ldots$. Eqs. (G.4)-(G.6) imply that Eq. (G.2) can be rewritten as

$$p^*_t = \frac{1}{2(1 + \alpha)} \left[p^*_{t-1} + p^*_{t+1|t}\right] + \frac{\alpha}{(1 + \alpha)} \left[A(L) + \hat{A}(L) + \frac{1}{1 + \alpha} \left[ B(L) + \hat{B}(L) - (1 + \rho_y)\right] y^N_t\right].$$  \hfill (G.7)

Similar to the expected money supply, rational expectation and AR(1) distributed shocks implies $p^*_{t+1|t} = \hat{F}(L)u_t + \hat{G}(L)y^N_t$, where $\hat{F}(L) = (f_{0\rho_y} + f_1) + f_2 L + f_3 L^2 + \ldots$ and $\hat{G}(L) = (g_{0\rho_y} + g_1) + g_2 L + g_3 L^2 + \ldots$.

Therefore,

$$F(L)u_t + G(L)y^N_t = \frac{1}{2(1 + \alpha)} \left[F(L)L^{-1} + \hat{F}(L)\right] u_t + \left[G(L)L^{-1} + \hat{G}(L)\right] y^N_t.$$  \hfill (G.8)
\[ + \frac{\alpha}{(1 + \alpha)} \left\{ A(L) + \tilde{A}(L) + \frac{1}{\alpha} (1 + \rho_u) \right\} u_t + \left[ B(L) + \tilde{B}(L) - (1 + \rho_y) \right] \psi_t^N \]  

Matching coefficients yields

\[ f_0 = \varphi (f_0 \rho_u + f_1) + \left( \frac{1}{2} - \varphi \right) \left[ \left( a_0 + \frac{1}{\alpha} \right) (1 + \rho_u) + a_1 \right] \]  

\[ f_j = \varphi (f_{j-1} + f_{j+1}) + \left( \frac{1}{2} - \varphi \right) (a_j + a_{j+1}), \quad \text{for } j = 1, 2, 3 \ldots \]  

\[ g_0 = \varphi (g_0 \rho_y + g_1) + \left( \frac{1}{2} - \varphi \right) \left[ (b_0 - 1)(1 + \rho_y) + b_1 \right] \]  

\[ g_j = \varphi (g_{j-1} + g_{j+1}) + \left( \frac{1}{2} - \varphi \right) (b_j + b_{j+1}), \quad \text{for } j = 1, 2, 3 \ldots \]  

where

\[ \varphi = \frac{1 - \alpha}{2(1 + \alpha)}. \]

References

Adam, K., 2007. Optimal monetary policy with imperfect common knowledge. J. Monetary Econ. 54, 267–301.


Ball, L., Mankiw, N.G., Reis, R., 2005. Monetary policy for inattentive economies. J. Monetary Econ. 52, 703–725.


