Review of Economic Studies (2014) **81**, 356–388 doi: 10.1093/restud/rdt024

Advance access publication 31 July 2013

# Exogenous Information, Endogenous Information, and Optimal Monetary Policy

LUIGI PACIELLO
*Einaudi Institute for Economics and Finance*
and
MIRKO WIEDERHOLT
*Goethe University Frankfurt*

*First version received February* 2011; *final version accepted March* 2013 (*Eds.*)

This article studies optimal monetary policy when decision-makers in firms choose how much attention they devote to aggregate conditions. When the amount of attention that decision-makers in firms devote to aggregate conditions is exogenous, complete price stabilization is optimal only in response to shocks that cause efficient fluctuations under perfect information. When decision-makers in firms choose how much attention they devote to aggregate conditions, complete price stabilization is optimal also in response to shocks that cause inefficient fluctuations under perfect information. Hence, recognizing that decision-makers in firms can choose how much attention they devote to aggregate conditions has major implications for optimal policy.

*Key words*: Optimal monetary policy, Dispersed information, Rational inattention, Information frictions

*JEL Codes*: E3, E5, D8.

## 1. INTRODUCTION

Decision-makers in firms have a limited amount of attention and they can choose how much attention they devote to aggregate conditions. What are the implications for optimal economic policy? What are the implications for optimal monetary policy?

To address these questions formally, we derive optimal monetary policy under two alternative assumptions. Initially, we assume that the amount of attention that decision-makers in firms devote to aggregate conditions is exogenous. Subsequently, we allow decision-makers in firms to choose how much attention they devote to aggregate conditions. With an exogenous allocation of attention by decision-makers in firms, complete price stabilization is the optimal monetary policy *only* in response to shocks that cause *efficient* fluctuations under perfect information. In contrast, when decision-makers in firms can choose how much attention they devote to aggregate conditions, complete price stabilization is the optimal monetary policy *also* in response to shocks that cause *inefficient* fluctuations under perfect information. The second result is the main result of this article. It turns out to hold in a wide range of models. The exact circumstances are described below.

There is a large literature on optimal monetary policy. Most of this literature studies optimal monetary policy in the New Keynesian framework (see Woodford (2003) or Gali (2008) for a detailed summary of the results). To make our results comparable to this benchmark in the literature on optimal monetary policy, we maintain several assumptions of the standard New Keynesian model: We assume that there is a large number of ex-ante identical firms supplying differentiated products and setting prices for these products; the monetary policy instrument is a nominal variable; and the central bank can affect real variables with the nominal monetary policy instrument because prices adjust slowly. The economy is subject to different types of shocks. The main policy question is how the central bank should adjust the monetary policy instrument in response to these shocks. We make two changes to the standard New Keynesian model. First, slow adjustment of prices to changes in aggregate conditions is due to limited attention by decision-makers in firms who set prices (henceforth "price setters") rather than price stickiness à la Calvo (1983). Second, once we allow price setters to choose how much attention they devote to aggregate conditions, the degree of price stickiness is no longer invariant to policy.

Following the literature on rational inattention (see Sims (2003)), we assume that price setters can in principle pay attention to any freely available information, but paying attention is costly. We model paying attention to aggregate conditions as receiving a noisy signal concerning aggregate conditions. A price setter who pays more attention receives a more precise signal.[1] Following Sims (2003), the amount of attention is quantified by uncertainty reduction. The cost of paying attention to aggregate conditions is assumed to be strictly increasing, smooth, and convex in the amount of attention devoted to aggregate conditions. The most common assumption in the rational inattention literature is that this cost function is linear. We allow for more general cost functions. The main prediction concerning the attention decision is that price setters pay more attention to aggregate conditions when the benefit of paying attention is higher.

We study optimal monetary policy under commitment assuming that the central bank aims to maximize expected utility of the representative household. We first derive the optimal monetary policy response to aggregate technology shocks. The key property of aggregate technology shocks in the benchmark model setup is that the response of the economy to aggregate technology shocks is efficient under perfect information and flexible prices. Furthermore, the central bank can replicate the perfect-information, flex-price response of the economy to aggregate technology shocks by stabilizing completely the profit-maximizing price in response to aggregate technology shocks. Moreover, when the profit-maximizing price does not move actual prices do not move. Hence, complete price stabilization in response to aggregate technology shocks is optimal. This argument applies in the model with exogenous attention by price setters, in the model with endogenous attention by price setters, and in the standard New Keynesian model. To the best of our knowledge, this argument is the only argument for complete price stabilization that has been proposed so far in the New Keynesian literature or in the literature on monetary models with information frictions. This argument applies to other shocks that cause efficient fluctuations under perfect information and flexible prices, so long as the central bank can replicate the perfect-information, flex-price allocation by stabilizing completely the profit-maximizing price.

We then derive the optimal monetary policy response to markup shocks. A markup shock is a shock to the elasticity of substitution between goods. The key property of markup shocks in the benchmark model setup is that the response of the economy to markup shocks is inefficient under perfect information and flexible prices. In the model with exogenous attention by price setters, complete price stabilization in response to markup shocks is never optimal. In contrast, in the model with endogenous attention by price setters, complete price stabilization in response

1. Think of the noise in the signal as the noise in the answers you get when you ask a sample of economists what the official CPI inflation rate for the U.S. was last year.

to markup shocks is optimal if the cost function for attention $f(\kappa)$ satisfies the condition $f''(\kappa)/f'(\kappa) \leq 2\ln(2)$. Here $\kappa$ denotes the attention devoted to aggregate conditions, $f'(\kappa)$ denotes the first derivative, and $f''(\kappa)$ denotes the second derivative. In sum, this article provides a new argument for complete price stabilization to be optimal. Furthermore, the level of the cost of paying attention does not matter: Multiplying the cost function for attention by an arbitrarily small positive constant does not affect whether the condition is satisfied. Moreover, this result for markup shocks extends to other shocks causing inefficient fluctuations under perfect information and flexible prices.

The intuition for these results is the following. The central bank could again replicate the perfect-information, flex-price response of the economy to markup shocks by completely stabilizing the profit-maximizing price in response to markup shocks, but since markup shocks cause inefficient fluctuations, this is not desirable. At the other extreme, note what happens when the central bank does not respond at all with the policy instrument to markup shocks. A positive markup shock increases the profit-maximizing price. Price setters put a positive weight on their noisy signals concerning the desired markup which causes inefficient price dispersion. Furthermore, the price level increases which given the constant monetary policy instrument causes a fall in consumption. To reduce inefficient price dispersion, the central bank can conduct a contractionary monetary policy after positive markup shocks. The profit-maximizing price then increases by less in response to positive markup shocks. Price setters put a smaller weight on their noisy signals concerning the desired markup which reduces inefficient price dispersion. Unfortunately, the contractionary monetary policy amplifies the fall in consumption. There exists a trade-off between inefficient price dispersion and inefficient consumption variance. Moreover, the marginal benefit of reducing inefficient price dispersion goes to zero as inefficient price dispersion goes to zero. Hence, in the model with exogenous attention by price setters, complete price stabilization in response to markup shocks is never optimal. The existence of a trade-off between inefficient price dispersion and inefficient consumption variance in the presence of shocks that cause inefficient fluctuations is emphasized a lot in the literature on optimal monetary policy. The result that complete price stabilization is suboptimal in response to these shocks is a classic result in monetary economics.

However, in the model with endogenous attention by price setters, there is one additional effect. Recall that the central bank reduces inefficient price dispersion by stabilizing the profit-maximizing price in response to markup shocks. The smaller the response of the profit-maximizing price to markup shocks, the smaller the incentive of price setters to pay attention to markup shocks. Hence, the exact same policy that reduces inefficient price dispersion also reduces price setters' incentive to pay attention to markup shocks. Next, consider consumption. A monetary policy of stabilizing the profit-maximizing price in response to markup shocks has two effects on consumption: First, a contractionary monetary policy after a positive markup shock has a direct negative effect on consumption. This effect is also present in the model with exogenous attention by price setters. Second, when price setters pay less attention to markup shocks, the price level increases by less after a positive markup shock and thus consumption falls by less after a positive markup shock. This effect is only present in the model with endogenous attention by price setters. When the cost function for attention satisfies the condition stated above, the new effect dominates. A monetary policy that reduces the response of the profit-maximizing price to markup shocks then reduces both inefficient price dispersion and inefficient consumption variance. The trade-off between inefficient price dispersion and inefficient consumption variance disappears. The optimal monetary policy is to reduce the response of the profit-maximizing price to markup shocks until price setters just pay no attention to markup shocks. Hence, actual prices do not respond to markup shocks. Complete price stabilization in response to markup shocks is optimal. Note that the optimality of complete price stabilization is a symptom of the fact that it is optimal to make price setters pay no attention to shocks that cause inefficient fluctuations. Having a sticky price level in response to these shocks is good not bad.

Given that the official primary goal of so many central banks is to maintain price stability, the insight that complete price stabilization can be optimal also in response to shocks causing inefficient fluctuations under perfect information and flexible prices seems important. In addition, the previous result yields an important insight about optimal central bank communication: If the cost function for attention satisfies the condition stated on the previous page and the central bank sets the policy instrument optimally, providing price setters with *easier access* to information about shocks that cause inefficient fluctuations *reduces* welfare. The reason is simple. The optimal policy is to discourage price setters from paying attention to these shocks. If the central bank makes it easier to pay attention to these shocks, the central bank has to reduce the response of the profit-maximizing price to these shocks even more to make price setters pay no attention to these shocks. For the same reason, making it more *difficult* to attend to these shocks *increases* welfare. In fact, if the cost of paying attention to markup shocks is sufficiently high, the central bank can attain the economy's efficient response to markup shocks. In particular, if the cost of paying attention to markup shocks is so high that the central bank no longer needs to actively discourage price setters from attending to markup shocks, the economy exhibits the first-best response to markup shocks: no response.

In the section on robustness and extensions, we relax many assumptions of the benchmark model setup. Initially, we assume that information processing noise is idiosyncratic. Thereafter we show that our results also hold when information processing noise is correlated across price setters.[2] Initially, we assume that price setters receive independent signals about aggregate technology and the desired markup and choose the precision of those signals. Subsequently, we show that optimal monetary policy in the model with endogenous information is the same when price setters can decide to receive signals concerning any linear combination of aggregate technology and the desired markup (*e.g*. price setters can pay attention to endogenous variables). Furthermore, we show that the results for aggregate technology shocks extend to other shocks causing efficient fluctuations under perfect information, while the results for markup shocks extend to other shocks causing inefficient fluctuations under perfect information. In addition, we study the effects of introducing sticky wages and imperfect information by the central bank into the benchmark model setup. In the model with exogenous information, each of these assumptions implies that complete price stabilization in response to aggregate technology shocks becomes suboptimal. Similarly, in the standard New Keynesian model, each of these assumptions implies that complete price stabilization in response to aggregate technology shocks becomes suboptimal. However, in the model with endogenous information, complete price stabilization in response to aggregate technology shocks *and* complete price stabilization in response to markup shocks *remain* optimal for a range of parameter values. Hence, the new argument for optimality of complete price stabilization presented in this article is more robust to variations of the model setup than the traditional argument that is based on the assumption that the central bank can replicate an efficient flexible-price, perfect-information solution.

This article is related to four recent papers on optimal monetary policy when price setters have imperfect information. Adam (2007) studies optimal monetary policy in an economy where price setters pay limited attention to aggregate conditions, but the amount of attention that price setters devote to aggregate conditions is exogenous. He shows that complete price stabilization is optimal in response to labour supply shocks but not in response to markup shocks. Ball *et al.* (2005) study optimal monetary policy in an economy where price setters update their information

2. We allow for an arbitrary degree of correlation. When noise is perfectly correlated across price setters, noise-driven consumption variance plays the role of price dispersion in the arguments given above.

sets with an exogenous probability. They show that complete price stabilization is optimal in response to aggregate technology shocks but not in response to markup shocks. Angeletos and La'O (2012) study optimal policy when price setters have imperfect information and the same individuals who set prices also take input decisions. They show that complete price stabilization is no longer optimal in response to aggregate technology shocks. They make this point in a model with an exogenous information structure. We revisit this point in our model with an endogenous information structure. We find that complete price stabilization is optimal both in response to aggregate technology shocks and in response to markup shocks for a range of parameter values. Lorenzoni (2010) studies optimal monetary policy in an economy where agents in the private sector receive noisy private and public signals about aggregate technology and the central bank cannot tell apart aggregate shocks to fundamentals (technology shocks) and aggregate shocks to beliefs (noise in the public signal). The information structure is again exogenous.

This article is also related to the literature on rational inattention. See, for example, Sims (2003, 2006, 2011), Luo (2008), Maćkowiak and Wiederholt (2009, 2013), Van Nieuwerburgh and Veldkamp (2009, 2010), Woodford (2009), Mondria (2010), Matejka (2011), and Paciello (2012). However, none of these papers studies optimal policy.

The rest of the article is organized as follows. Section 2 describes the benchmark model setup. Section 3 states the problem of the central bank in the model with exogenous information and in the model with endogenous information. Section 4 presents the optimal monetary policy response to aggregate technology shocks. Section 5 derives the optimal monetary policy response to markup shocks. Section 6 contains all the additional results described above. Section 7 concludes.

## 2. MODEL SETUP

The economy is populated by firms, a representative household, and a government.

### 2.1. *Household*

The household's preferences in period zero over sequences of consumption and labour supply $\{C_t, L_t\}_{t=0}^{\infty}$ are given by

$$E_0\left[\sum_{t=0}^{\infty}\beta^t\left(\frac{C_t^{1-\gamma}-1}{1-\gamma}-\frac{L_t^{1+\psi}}{1+\psi}\right)\right], \tag{1}$$

where $C_t$ is composite consumption and $L_t$ is labour supply in period $t$. The parameter $\beta \in (0,1)$ is the discount factor, $\gamma > 0$ is the inverse of the intertemporal elasticity of substitution, $\psi \geq 0$ is the inverse of the Frisch elasticity of labour supply, and $E_0$ denotes the expectation operator conditioned on the household's information in period zero. Composite consumption is given by a Dixit–Stiglitz aggregator

$$C_t = \left(\frac{1}{I}\sum_{i=1}^{I} C_{i,t}^{\frac{1}{1+\Lambda_t}}\right)^{1+\Lambda_t}, \tag{2}$$

where $C_{i,t}$ is consumption of good $i$ in period $t$. There are $I$ different consumption goods and the elasticity of substitution between those different consumption goods in period $t$ equals $(1+1/\Lambda_t)$. We call the variable $\Lambda_t$ the desired markup because it equals the desired markup by firms in period $t$. The log of the desired markup follows a stationary Gaussian first-order autoregressive process

$$\ln(\Lambda_t) = (1-\rho_\lambda)\ln(\Lambda) + \rho_\lambda \ln\left(\Lambda_{t-1}\right) + \nu_t, \tag{3}$$

where the parameter $\Lambda > 0$, the parameter $\rho_\lambda \in [0,1)$, and the innovation $\nu_t$ is $i.i.d.N\left(0, \sigma_\nu^2\right)$.

The flow budget constraint of the representative household in period $t$ reads

$$\sum_{i=1}^{I} P_{i,t} C_{i,t} + B_t = R_{t-1} B_{t-1} + W_t L_t + D_t - T_t. \tag{4}$$

Here $P_{i,t}$ is the price of good $i$ in period $t$, $B_t$ are the representative household's holdings of government bonds between period $t$ and period $t+1$, $R_t$ is the nominal gross interest rate on those bond holdings, $W_t$ is the nominal wage rate, $D_t$ are nominal aggregate profits, and $T_t$ are nominal lump sum taxes in period $t$. The representative household also faces a no-Ponzi-scheme condition.

In every period, the representative household chooses a consumption vector, labour supply, and bond holdings. The household takes as given the prices of consumption goods, the wage rate, the interest rate, aggregate profits, and lump sum taxes.

### 2.2. *Firms*

There are $I$ firms. Firm $i$ supplies good $i$. The technology of firm $i$ is given by

$$Y_{i,t} = A_t L_{i,t}^{\alpha}, \tag{5}$$

where $Y_{i,t}$ is output and $L_{i,t}$ is labour input of firm $i$ in period $t$. $A_t$ is aggregate productivity in period $t$. The parameter $\alpha \in (0,1]$ is the elasticity of output with respect to labour input. The log of aggregate productivity follows a stationary Gaussian first-order autoregressive process

$$\ln(A_t) = \rho_a \ln\left(A_{t-1}\right) + \varepsilon_t, \tag{6}$$

where the parameter $\rho_a \in [0,1)$ and the innovation $\varepsilon_t$ is $i.i.d.N\left(0, \sigma_\varepsilon^2\right)$. The processes $\{A_t\}_{t=0}^{\infty}$ and $\{\Lambda_t\}_{t=0}^{\infty}$ are independent.

Nominal profit of firm $i$ in period $t$ equals

$$\left(1+\tau_p\right) P_{i,t} Y_{i,t} - W_t L_{i,t}, \tag{7}$$

where $\tau_p$ is a production subsidy paid by the government.

In every period, each firm sets a price and commits to supply any quantity at that price. Each firm takes as given the nominal wage rate, composite consumption, and the following price index

$$P_t = \left( \frac{1}{I} \sum_{i=1}^{I} P_{i,t}^{-\frac{1}{\Lambda_t}} \right)^{-\Lambda_t} I. \tag{8}$$

### 2.3. *Government*

There is a monetary authority and a fiscal authority. The monetary authority controls directly nominal spending and commits to the following rule for nominal spending

$$\ln(M_t) = F_t(L)\varepsilon_t + G_t(L)\nu_t, \tag{9}$$

where $M_t \equiv \sum_{i=1}^{I} P_{i,t} C_{i,t}$ denotes nominal spending in period $t$. $F_t(L)$ and $G_t(L)$ are infinite-order lag polynomials which may depend on $t$. The last equation simply says that the log of nominal spending in period $t$ can be any linear function of the sequence of shocks up to and including period $t$. We ask the question which linear function is optimal.

To derive optimal monetary policy, we initially assume that the central bank controls directly nominal spending and has perfect information. In Section 6, we then show that the set of equilibria that the central bank can implement with a nominal spending rule of the form (9) equals the set of equilibria that the central bank can implement with an interest rate rule of the form

$$\ln(R_t) = F_t(L)\varepsilon_t + G_t(L)\nu_t. \tag{10}$$

We begin with a nominal spending rule rather than an interest rate rule because it makes the results about optimal monetary policy more transparent and multiplicity of equilibria at a given monetary policy arises less easily. We postpone the discussion of unique implementation in the case of an interest rate rule to Section 6. In Section 6, we also study optimal monetary policy when the central bank has imperfect information.

Next, consider fiscal policy. The government budget constraint in period $t$ reads

$$T_t + B_t = R_{t-1}B_{t-1} + \tau_p\left(\sum_{i=1}^{I} P_{i,t}Y_{i,t}\right). \tag{11}$$

The government has to finance maturing nominal government bonds and the production subsidy. The government can collect lump sum taxes or issue new one-period nominal government bonds. The fiscal authority sets the production subsidy so as to correct the distortion arising from monopolistic competition in the non-stochastic steady state:

$$\tau_p = \Lambda. \tag{12}$$

We assume that monetary policy is active and fiscal policy is passive in the sense of Leeper (1991).

### 2.4. *Information*

The date $t$ information set of the decision-maker who sets the price of good $i$ is

$$\mathcal{I}_{i,t} = \mathcal{I}_{i,-1} \cup \left\{s_{i,0}, s_{i,1}, \ldots, s_{i,t}\right\}, \tag{13}$$

where $\mathcal{I}_{i,-1}$ is the initial information set of the price setter in firm $i$ and $s_{i,t}$ is the signal that he or she receives in period $t$. This signal is a two-dimensional vector consisting of a noisy signal about aggregate technology and a noisy signal about the desired markup

$$s_{i,t} = \begin{pmatrix} \ln(A_t) + \eta_{i,t} \\ \ln(\Lambda_t/\Lambda) + \zeta_{i,t} \end{pmatrix}. \tag{14}$$

We assume that the noise is due to limited attention and has the following properties: (i) the processes $\left\{\eta_{i,t}\right\}_{t=0}^{\infty}$ and $\left\{\zeta_{i,t}\right\}_{t=0}^{\infty}$ are independent of the processes $\{A_t\}_{t=0}^{\infty}$ and $\{\Lambda_t\}_{t=0}^{\infty}$, (ii) the processes $\left\{\eta_{i,t}\right\}_{t=0}^{\infty}$ and $\left\{\zeta_{i,t}\right\}_{t=0}^{\infty}$ are independent across firms and independent of each other, and (iii) $\eta_{i,t}$ and $\zeta_{i,t}$ follow Gaussian white noise processes with variances $\sigma_\eta^2$ and $\sigma_\zeta^2$. The assumption that price setters receive separate signals about aggregate technology and the desired markup is for ease of exposition. The assumption that noise is independent across firms accords well with the idea that the source of noise in signals is limited attention by individual decision-makers rather than lack of publicly available information. The reader is reminded that both assumptions are relaxed in Section 6.

We consider two models. In the first model, the attention that price setters devote to aggregate conditions is exogenous (*i.e.* $1/\sigma_\eta^2$ and $1/\sigma_\zeta^2$ are exogenous). In the second model, price setters choose how much attention they devote to aggregate conditions (*i.e.* $1/\sigma_\eta^2$ and $1/\sigma_\zeta^2$ are endogenous). The timing is as follows. In period minus one, price setters choose how much attention they devote to aggregate conditions so as to maximize expected profits net of the cost of paying attention. In the following periods, price setters receive signals with the chosen precisions and take optimal pricing decisions given these signals. Formally, the price setter in firm $i$ solves the following decision problem in period minus one

$$\max_{\left(1/\sigma_\eta^2, 1/\sigma_\zeta^2\right)\in\mathbb{R}_+^2} \left\{ E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t \pi\left(P_{i,t}, P_t, C_t, W_t, A_t, \Lambda_t\right)\right] - \frac{1}{1-\beta} f(\kappa)\right\}, \tag{15}$$

subject to

$$P_{i,t} = \arg\max_{X\in\mathbb{R}_{++}} E[\pi(X, P_t, C_t, W_t, A_t, \Lambda_t)|\mathcal{I}_{i,t}], \tag{16}$$

and

$$\kappa = h\left(A_t, \Lambda_t|\mathcal{I}_{i,t-1}\right) - h\left(A_t, \Lambda_t|\mathcal{I}_{i,t}\right), \tag{17}$$

where the information set $\mathcal{I}_{i,t}$ is given by equations (13)–(14). Consider first objective (15). Here $E_{i,-1}$ denotes the expectation operator conditioned on the information of the price setter in firm $i$ in period minus one. The function $\pi$ denotes the real profit function defined as the nominal profit function divided by $P_t$ times the marginal utility of consumption of the representative household. The variable $\kappa$ is the amount of attention that the price setter in firm $i$ devotes to aggregate conditions. Following Sims (2003), we quantify the amount of attention devoted to aggregate conditions by uncertainty reduction, where uncertainty is measured by entropy. Formally, $h\left(A_t, \Lambda_t|\mathcal{I}_{i,t-1}\right)$ denotes the conditional entropy of $A_t$ and $\Lambda_t$ given $\mathcal{I}_{i,t-1}$, whereas $h\left(A_t, \Lambda_t|\mathcal{I}_{i,t}\right)$ denotes the conditional entropy of $A_t$ and $\Lambda_t$ given $\mathcal{I}_{i,t}$. The difference between the two is the uncertainty reduction due to information received in period $t$. The term $f(\kappa)$ is the per-period cost of paying attention to aggregate conditions. The function $f:\mathbb{R}\to\mathbb{R}$ is assumed to be twice continuously differentiable on its domain and non-negative, strictly increasing, and convex on $\mathbb{R}_+$.[3] Finally, equation (16) specifies the price setting behaviour for a given realization of the signals. The basic trade-off is that paying more attention to aggregate conditions improves pricing decisions but is costly.

Finally, we make a simplifying assumption. To abstract from transitional dynamics in conditional second moments, we assume that at the end of period minus one (*i.e.* after decision-makers have chosen signal precisions), the decision-makers receive information such that: (i) the conditional distribution of $(\ln(A_0), \ln(\Lambda_0))$ given information at the end of period minus one is normal, and (ii) the conditional covariance matrix of $(\ln(A_0), \ln(\Lambda_0))$ given information at the end of period minus one equals the steady-state conditional covariance matrix of $(\ln(A_t), \ln(\Lambda_t))$ given information in period $t-1$.

3. Examples include $f(\kappa) = \mu\kappa$ and $f(\kappa) = \mu 2^{2\kappa}$ with $\mu > 0$. The rational inattention literature has so far almost exclusively used these two cost functions. The second cost function arises under certain assumptions when the cost of paying more attention to something is the opportunity cost of paying less attention to something else. See, for example, Maćkowiak and Wiederholt (2009).

Note that in the decision problem (15)–(17) price setters choose constant signal precisions once and for all. All propositions in this article also hold when they choose signal precisions as a function of time in period minus one or when they choose signal precisions period by period.

We assume that the representative household has perfect information. We make this assumption for two reasons. First, this assumption facilitates the comparison of our results to the results about optimal monetary policy in the basic New Keynesian model, where the only friction apart from monopolistic competition is price stickiness. Second, this assumption allows us to isolate the implications of limited attention by price setters in firms for optimal monetary policy.

### 2.5. *Aggregation*

When computing the price index, terms will appear that are linear in $\frac{1}{I}\sum_{i=1}^{I}\eta_{i,t}$ and $\frac{1}{I}\sum_{i=1}^{I}\zeta_{i,t}$. These averages are random variables with mean zero and variance $\frac{1}{I}\sigma_\eta^2$ and $\frac{1}{I}\sigma_\zeta^2$, respectively. We will neglect these terms because these terms have mean zero and a variance that can be made arbitrarily small by setting the number of firms sufficiently high. For example, one can set $I=10^{100}$. Alternatively, one could work with a continuum of firms and apply the law of large numbers in Uhlig (1995). We work with a finite number of firms rather than a continuum of firms because we find that it makes the derivation of the central bank's objective in the next section more transparent.

## 3. THE RAMSEY PROBLEM

In this section, we state the optimization problem of the central bank. The central bank wants to commit to the policy rule that maximizes expected utility of the representative household. In the literature on optimal monetary policy, it is common practice to study the central bank's optimization problem after a log-quadratic approximation of the central bank's objective and a log-linear approximation of the equilibrium conditions. See, for example, Woodford (2003), Gali (2008), Ball *et al.* (2005), and Adam (2007). We follow this common practice. This makes our results comparable to their results.

### 3.1. *Derivation of the central bank's objective*

One can derive a simple expression for period utility at a feasible allocation by substituting the technology and the consumption aggregator into the period utility function. At a feasible allocation, the representative household has to supply the labour that is needed to produce the consumption vector

$$L_t=\sum_{i=1}^{I}\left(\frac{C_{i,t}}{A_t}\right)^{\frac{1}{\alpha}}.$$

Furthermore, equation (2) for the consumption aggregator can be written as

$$\hat{C}_{I,t}=\left(I-\sum_{i=1}^{I-1}\hat{C}_{i,t}^{\frac{1}{1+\Lambda_t}}\right)^{1+\Lambda_t},$$

where $\hat{C}_{i,t}=\left(C_{i,t}/C_t\right)$ denotes relative consumption of good $i$ in period $t$. Substituting the last two equations into the period utility function in (1) yields the following expression for period utility at a feasible allocation

$$U\left(C_t,\hat{C}_{1,t},\ldots,\hat{C}_{I-1,t},A_t,\Lambda_t\right)=\frac{C_t^{1-\gamma}-1}{1-\gamma}-\frac{1}{1+\psi}\left(\frac{C_t}{A_t}\right)^{\frac{1}{\alpha}(1+\psi)}$$
$$\times\left[\sum_{i=1}^{I-1}\hat{C}_{i,t}^{\frac{1}{\alpha}}+\left(I-\sum_{i=1}^{I-1}\hat{C}_{i,t}^{\frac{1}{1+\Lambda_t}}\right)^{\frac{1}{\alpha}(1+\Lambda_t)}\right]^{1+\psi}. \qquad (18)$$

Period utility at a feasible allocation is a function only of the consumption vector, aggregate productivity, and the desired markup in that period.

The efficient allocation in period $t$ is defined as the feasible allocation in period $t$ that maximizes utility of the representative household. Maximizing expression (18) yields

$$C_t^*=\left(\frac{\alpha}{I^{1+\psi}}\right)^{\frac{1}{\gamma-1+\frac{1}{\alpha}(1+\psi)}}A_t^{\frac{\frac{1}{\alpha}(1+\psi)}{\gamma-1+\frac{1}{\alpha}(1+\psi)}},$$

and, for all $i=1,\ldots,I-1$,

$$\hat{C}_{i,t}^*=1.$$

The efficient consumption level in period $t$ is strictly increasing in aggregate productivity in that period and is independent of the desired markup, while the efficient consumption mix in period $t$ is to consume an equal amount of each good.

We will work with a log-quadratic approximation to the period utility function (18) around the non-stochastic steady state. From now on, variables without time subscript denote values in the non-stochastic steady state and small letters denote log-deviations from the non-stochastic steady state (*e.g.* $c_t=\ln(C_t/C)$). Expressing the period utility function $U$ defined by equation (18) in terms of log-deviations from the non-stochastic steady state and using the fact that the non-stochastic steady state is efficient due to the production subsidy (12) (*i.e.* $C=C^*$ and $\hat{C}_i=1$) yields the following expression for period utility at a feasible allocation in terms of log-deviations

$$u\left(c_t,\hat{c}_{1,t},\ldots,\hat{c}_{I-1,t},a_t,\lambda_t\right)=\frac{C^{1-\gamma}e^{(1-\gamma)c_t}-1}{1-\gamma}$$
$$-\frac{C^{1-\gamma}e^{\frac{1}{\alpha}(1+\psi)(c_t-a_t)}}{\frac{1}{\alpha}(1+\psi)}\left[\frac{1}{I}\sum_{i=1}^{I-1}e^{\frac{1}{\alpha}\hat{c}_{i,t}}+\frac{1}{I}\left(I-\sum_{i=1}^{I-1}e^{\hat{c}_{i,t}\frac{1}{1+\Lambda e^{\lambda_t}}}\right)^{\frac{1}{\alpha}\left(1+\Lambda e^{\lambda_t}\right)}\right]^{1+\psi}. \qquad (19)$$

Let $\tilde{u}$ denote the second-order Taylor approximation to this period utility function at the origin. Let $E$ denote the unconditional expectation operator. Let $x_t$ denote the vector of endogenous variables and let $z_t$ denote the vector of exogenous variables (*i.e.* $x_t'=\left(c_t\ \hat{c}_{1,t}\ \cdots\ \hat{c}_{I-1,t}\right)$ and $z_t'=\left(a_t\ \lambda_t\right)$). Under a regularity condition that is stated in Supplementary Appendix A and that is always satisfied in the models that we consider, we have

$$E\left[\sum_{t=0}^{\infty}\beta^t\tilde{u}(x_t,z_t)\right]-E\left[\sum_{t=0}^{\infty}\beta^t\tilde{u}\left(x_t^*,z_t\right)\right]=\sum_{t=0}^{\infty}\beta^tE\left[\frac{1}{2}\left(x_t-x_t^*\right)'H\left(x_t-x_t^*\right)\right], \qquad (20)$$

where the matrix $H$ is given by

$$H=-C^{1-\gamma}\begin{bmatrix} \gamma-1+\frac{1}{\alpha}(1+\psi) & 0 & \cdots & \cdots & 0 \\ 0 & 2\frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \cdots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} \\ \vdots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} \\ 0 & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \cdots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & 2\frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} \end{bmatrix},$$

and the efficient vector of endogenous variables $x_t^*$ is given by

$$c_t^*=\frac{\frac{1}{\alpha}(1+\psi)}{\gamma-1+\frac{1}{\alpha}(1+\psi)}a_t,$$

and

$$\hat{c}_{i,t}^*=0.$$

In words, after the log-quadratic approximation to the period utility function (18) around the non-stochastic steady state, the loss in expected utility in the case of deviations of the actual consumption vector from the efficient consumption vector is given by equation (20). The upper-left element of the matrix $H$ determines the loss in utility in the case of an inefficient consumption level (aggregate inefficiency). The lower-right block of the matrix $H$ determines the loss in utility in the case of an inefficient consumption mix (cross-sectional inefficiency).

### 3.2. *The Ramsey problem*

The log-quadratic approximation of the central bank's objective and a log-linear approximation of the equilibrium conditions around the non-stochastic steady state yields the following Ramsey problem for the economy with an exogenous information structure:

$$\min_{\{F_t(L),G_t(L)\}_{t=0}^{\infty}}\sum_{t=0}^{\infty}\beta^t E\left[\left(c_t-c_t^*\right)^2+\delta\frac{1}{I}\sum_{i=1}^{I}\left(p_{i,t}-p_t\right)^2\right], \tag{21}$$

subject to

$$c_t^*=\frac{\phi_a}{\phi_c}a_t, \tag{22}$$

$$c_t=m_t-p_t, \tag{23}$$

$$p_t=\frac{1}{I}\sum_{i=1}^{I}p_{i,t}, \tag{24}$$

$$p_{i,t}=E\left[p_{i,t}^*|\mathcal{I}_{i,t}\right], \tag{25}$$

$$p_{i,t}^*=p_t+\phi_c c_t-\phi_a a_t+\phi_\lambda\lambda_t, \tag{26}$$

$$\mathcal{I}_{i,t}=\mathcal{I}_{i,-1}\cup\left\{s_{i,0},s_{i,1},\ldots,s_{i,t}\right\}, \tag{27}$$

$$s_{i,t}=\begin{pmatrix}a_t+\eta_{i,t}\\ \lambda_t+\zeta_{i,t}\end{pmatrix}, \tag{28}$$

$$a_t=\rho_a a_{t-1}+\varepsilon_t, \tag{29}$$

$$\lambda_t=\rho_\lambda \lambda_{t-1}+\nu_t, \tag{30}$$

and

$$m_t=F_t(L)\varepsilon_t+G_t(L)\nu_t. \tag{31}$$

Here

$$\phi_c=\frac{\frac{\psi}{\alpha}+\gamma+\frac{1-\alpha}{\alpha}}{1+\frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}}>0,\quad \phi_a=\frac{\frac{\psi}{\alpha}+\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}}>0, \tag{32}$$

$$\phi_\lambda=\frac{\frac{\Lambda}{1+\Lambda}}{1+\frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}}>0,\quad \delta=\frac{\frac{1+\Lambda-\alpha}{(1+\Lambda)\alpha}\left(1+\frac{1}{\Lambda}\right)^2}{\gamma-1+\frac{1}{\alpha}(1+\psi)}>0. \tag{33}$$

The objective (21) follows from substituting the log-linearized demand function for good $i$ into equation (20) and using equation (24). The variable $c_t^*$ is the efficient consumption level in period $t$ and the coefficient $\delta$ is the relative weight on cross-sectional inefficiency versus aggregate inefficiency in the central bank's objective. The log-linear equation (23) follows from substituting the demand function for good $i$ into the definition of nominal spending. Equation (24) follows from log-linearizing the price index (8). Equation (25) states that each firm sets a price equal to the conditional expectation of the profit-maximizing price. The variable $p_{i,t}^*$ is the profit-maximizing price. Equations (27)–(31) simply restate assumptions of the model.

In the economy with an exogenous information structure, the amount of attention that price setters devote to aggregate conditions is exogenous (*i.e.* $1/\sigma_\eta^2$ and $1/\sigma_\zeta^2$ are exogenous). In the economy with an endogenous information structure, price setters choose how much attention they devote to aggregate conditions (*i.e.* $1/\sigma_\eta^2$ and $1/\sigma_\zeta^2$ are endogenous). After a log-quadratic approximation to the real profit function $\pi$ around the non-stochastic steady state, the attention problem (15)–(17) of the price setter in firm $i$ reads

$$\min_{\left(1/\sigma_\eta^2,1/\sigma_\zeta^2\right)\in\mathbb{R}_+^2}\left\{E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t\frac{\omega}{2}\left(p_{i,t}-p_{i,t}^*\right)^2\right]+\frac{1}{1-\beta}f(\kappa)\right\}, \tag{34}$$

subject to

$$p_{i,t}=E\left[p_{i,t}^*|\mathcal{I}_{i,t}\right], \tag{35}$$

and

$$\kappa=\frac{1}{2}\log_2\left(\frac{\sigma_{a|t-1}^2}{\sigma_{a|t}^2}\right)+\frac{1}{2}\log_2\left(\frac{\sigma_{\lambda|t-1}^2}{\sigma_{\lambda|t}^2}\right). \tag{36}$$

Here

$$\omega=C^{-\gamma}\frac{WL_i}{P}\frac{\frac{1+\Lambda}{\Lambda}}{\alpha}\left(1+\frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}\right). \tag{37}$$

The price setter aims to minimize the expected loss in profits due to deviations of the actual price from the profit-maximizing price plus the cost of paying attention. The price setter anticipates that in each future period he or she will set the price that equals the conditional expectation of the profit-maximizing price. The expression for information flow (17) reduces to equation (36) because the conditional distribution of $(a_t, \lambda_t)$ is Gaussian given $\mathcal{I}_{i,t-1}$ and $\mathcal{I}_{i,t}$ and because $a_t$ and $\lambda_t$ are conditionally independent given $\mathcal{I}_{i,t-1}$ and $\mathcal{I}_{i,t}$. Here $\sigma^2_{a|t-1}$ denotes the conditional variance of $a_t$ given $\mathcal{I}_{i,t-1}$, while $\sigma^2_{a|t}$ denotes the conditional variance of $a_t$ given $\mathcal{I}_{i,t}$. The term $L_i$ in (37) is the steady-state labour input of an individual firm. In the economy with an endogenous information structure, the optimization problem (34)–(37) is part of the constraints of the Ramsey problem.[4]

### 3.3. *Perfect-information allocation*

To understand optimal monetary policy, it will be useful to compare the equilibrium allocation under perfect information to the efficient allocation. When price setters have perfect information, each firm charges the profit-maximizing price and equations (24)–(26) imply $p_{i,t} = p^*_{i,t} = p_t$ and

$$c_t = \frac{\phi_a}{\phi_c} a_t - \frac{\phi_\lambda}{\phi_c} \lambda_t. \tag{38}$$

Comparing equation (38) to equation (22) and noting that inefficient price dispersion equals zero under perfect information clearly shows the key property of aggregate technology and markup shocks in the benchmark model setup: The perfect-information, flex-price response of the economy to aggregate technology shocks is efficient, while the perfect-information, flex-price response of the economy to markup shocks is inefficient.[5]

Moreover, when price setters have imperfect information, the central bank can still replicate the equilibrium allocation under perfect information. Substituting equation (23) into equation (26) yields the following expression for the profit-maximizing price

$$p^*_{i,t} = (1-\phi_c) p_t + \phi_c \left( m_t - \frac{\phi_a}{\phi_c} a_t + \frac{\phi_\lambda}{\phi_c} \lambda_t \right). \tag{39}$$

By setting $m_t = \frac{\phi_a}{\phi_c} a_t - \frac{\phi_\lambda}{\phi_c} \lambda_t$, the central bank can completely offset the effect of shocks on the profit-maximizing price. Equation (39) then reduces to $p^*_{i,t} = (1-\phi_c) p_t$ and the unique equilibrium is: $p_t = 0$, $c_t = m_t = \frac{\phi_a}{\phi_c} a_t - \frac{\phi_\lambda}{\phi_c} \lambda_t$, and $p_{i,t} - p_t = 0$. An immediate consequence is that welfare at the optimal monetary policy when price setters have imperfect information has to be at least as high as welfare when price setters have perfect information.

## 4. OPTIMAL MONETARY POLICY RESPONSE TO TECHNOLOGY SHOCKS

In this section, we derive the optimal monetary policy response to aggregate technology shocks when there are only technology shocks. In the next section, we derive the optimal monetary policy response to markup shocks when there are only markup shocks. In Section 6, we cover the case when there are both aggregate technology shocks and markup shocks.

4. We state equation (36) for the case when there are both technology shocks and markup shocks. If there are only technology shocks or only markup shocks, there is only one term on the right-hand side of equation (36).

5. The response of the economy to markup shocks under perfect information is inefficient because under perfect information firms vary the actual markup with the desired markup, which causes inefficient consumption fluctuations.

Assume that there are no markup shocks. By conducting the policy $m_t = \frac{\phi_a}{\phi_c} a_t$, the central bank can replicate the response of the economy to aggregate technology shocks under perfect information, and the perfect-information response of the economy to aggregate technology shocks is efficient. See the previous section. Hence, with this policy the central bank attains the efficient response of the economy to aggregate technology shocks. Furthermore, with any other policy the central bank does not attain the efficient response of the economy to aggregate technology shocks. Suppose $m_t \neq \frac{\phi_a}{\phi_c} a_t$. If price setters put weight on their noisy signals concerning aggregate technology there is inefficient price dispersion, and if price setters put no weight on their noisy signals concerning aggregate technology the response of consumption to aggregate technology shocks is inefficient. Moreover, when the central bank conducts the policy $m_t = \frac{\phi_a}{\phi_c} a_t$ the price level does not respond to aggregate technology shocks because the profit-maximizing price does not respond to those shocks. See the previous section. These results hold in the model with an exogenous information structure and the model with an endogenous information structure. We arrive at the following proposition.

**Proposition 1.** *Assume that $\sigma_\lambda^2 = \lambda_{-1} = 0$. Consider the Ramsey problem (21)–(33), where the variances of noise $\sigma_\eta^2$ and $\sigma_\zeta^2$ are exogenous. If $\sigma_\eta^2 > 0$, the unique optimal monetary policy response to aggregate technology shocks is*

$$F_t(L)\varepsilon_t = \frac{\phi_a}{\phi_c} a_t. \tag{40}$$

*At the optimal monetary policy, the price level does not respond to aggregate technology shocks. Next, consider the Ramsey problem (21)–(37), where the signal precisions $1/\sigma_\eta^2$ and $1/\sigma_\zeta^2$ are given by the solution to problem (34)–(37). The unique optimal monetary policy response to aggregate technology shocks is again given by equation (40). At the optimal monetary policy, the price level does not respond to aggregate technology shocks.*

## 5. OPTIMAL MONETARY POLICY RESPONSE TO MARKUP SHOCKS

We now derive the optimal monetary policy response to markup shocks. Our main result is that if the cost function for attention satisfies $f''(\kappa)/f'(\kappa) \leq 2\ln(2)$, the optimal monetary policy is qualitatively different in the two models: In the model with exogenous information, there is a trade-off between inefficient price dispersion and inefficient consumption variance, and complete price stabilization in response to markup shocks is never optimal; in the model with endogenous information, there is no trade-off between inefficient price dispersion and inefficient consumption variance, and complete price stabilization in response to markup shocks is always optimal. Moreover, in the latter model, the optimal monetary policy discourages price setters from paying attention to markup shocks and giving price setters easier access to aggregate information reduces welfare.

For ease of exposition, we assume in this section that there are no aggregate technology shocks. Sections 5.1 and 5.2 cover the case of an i.i.d. and autocorrelated desired markup, respectively. Section 6 covers the case of both technology shocks and markup shocks.

### 5.1. *White noise case*

In the model with an exogenous information structure, the optimal monetary policy response to markup shocks is given by the following proposition.

**Proposition 2.** *(Exogenous information structure) Consider the Ramsey problem (21)–(33), where the variances of noise $\sigma_\eta^2$ and $\sigma_\zeta^2$ are exogenous. Assume $\sigma_a^2 = a_{-1} = 0$, $\sigma_\lambda^2 > 0$, and $\rho_\lambda = 0$. Consider policies of the form $G_t(L)\nu_t = g_0\nu_t$ and equilibria of the form $p_t = \theta\lambda_t$. There exists a unique equilibrium for any policy $g_0 \in \mathbb{R}$. If $\sigma_\zeta^2 > 0$, the unique optimal monetary policy is*

$$g_0^* = \frac{(1-\delta\phi_c)\phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2} + \delta\phi_c^2}. \tag{41}$$

*At the optimal monetary policy, the price level strictly increases in response to a positive markup shock, there is inefficient price dispersion, and composite consumption strictly falls in response to a positive markup shock:*

$$\begin{aligned} p_{i,t} - p_t &= \frac{\phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2} + \delta\phi_c^2}\zeta_{i,t}, \\ p_t &= \frac{\phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2} + \delta\phi_c^2}\lambda_t, \\ c_t &= -\frac{\delta\phi_c\phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2} + \delta\phi_c^2}\lambda_t. \end{aligned}$$

*Proof* See Appendix A. ||

The main result in Proposition 2 is that in the model with an exogenous information structure complete price stabilization in response to markup shocks is *never* optimal. To understand this result, note first what happens when the central bank does not respond with the monetary policy instrument to markup shocks (*i.e.* $g_0 = 0$). At the inaction policy, a positive markup shock raises the profit-maximizing price. Price setters thus put a positive weight on their noisy signals about the desired markup which causes inefficient price dispersion. Furthermore, the price level increases which causes a fall in consumption. To reduce inefficient price dispersion, the central bank can lower the response of the profit-maximizing price to a positive markup shock with a contractionary monetary policy (*i.e.* $g_0 < 0$). The profit-maximizing price then increases by less after a positive markup shock. Price setters therefore put a smaller weight on their noisy signals about the desired markup which reduces inefficient price dispersion. See equations (39), (25), and (28). Unfortunately, the contractionary monetary policy amplifies the fall in consumption after a positive markup shock. There exists a trade-off between inefficient price dispersion and inefficient consumption variance: Decreasing one increases the other. Furthermore, driving inefficient price dispersion to zero (by stabilizing completely the profit-maximizing price and thereby prices) is never optimal because as inefficient price dispersion goes to zero the benefit of further reducing inefficient price dispersion goes to zero while the cost of further reducing inefficient price dispersion increases. Hence, complete price stabilization in response to markup shocks is never optimal.

The result that there is a trade-off between price dispersion and consumption variance in the presence of markup shocks and that complete price stabilization in response to these shocks is never optimal are classic results in monetary economics. They hold in a wide variety of other models. For the Calvo (1983) model of price stickiness, see Woodford (2003) and Gali (2008); for the sticky information model of price stickiness, see Ball *et al.* (2005); and for a noisy signal model of price stickiness, see Adam (2007). In all these models, the basic logic behind these two classic results is the one described above. A difference across models is the source of price dispersion: staggered price setting, staggered planning, and idiosyncratic noise. However, the basic logic that stabilizing the profit-maximizing price reduces price dispersion and increases consumption variance is the same across models.

Let us turn to the model with an endogenous information structure. In that model there is one additional effect. This effect changes optimal policy qualitatively. The new feature in the model with endogenous information is that the attention of price setters (and thus the degree of price stickiness) is no longer invariant to policy. Assume as before that $\sigma_a^2 = a_{-1} = 0$, $\sigma_\lambda^2 > 0$, and $\rho_\lambda = 0$. Consider policies of the form $G_t(L)\nu_t = g_0\nu_t$ and equilibria of the form $p_t = \theta\lambda_t$. Equation (39) implies that the profit-maximizing price of good $i$ in period $t$ equals

$$p_{i,t}^* = [(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]\lambda_t.$$

It follows from equations (34) to (36) that the attention problem of the price setter reads

$$\min_{\kappa\in\mathbb{R}_+} \frac{1}{1-\beta}\left[\frac{\omega}{2}\sigma_{p^*}^2 2^{-2\kappa} + f(\kappa)\right], \tag{42}$$

where $\sigma_{p^*}^2$ denotes the variance of the profit-maximizing price due to markup shocks and $\kappa$ equals the attention devoted to markup shocks. The first term in brackets reflects the benefit of paying attention. The second term in brackets is the cost of paying attention. Recall from the previous paragraphs that the central bank reduces inefficient price dispersion by reducing the response of the profit-maximizing price to markup shocks. The smaller the response of the profit-maximizing price to markup shocks, the smaller price setters' incentive to pay attention to markup shocks. *Hence, the exact same policy that reduces inefficient price dispersion also reduces price setters' incentive to pay attention to markup shocks.*

We now characterize the set of equilibria for a given monetary policy in the model with endogenous information. The price of good $i$ in period $t$ equals

$$p_{i,t} = [(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]E\left[\lambda_t|\mathcal{I}_{i,t}\right],$$

where

$$E\left[\lambda_t|\mathcal{I}_{i,t}\right] = \frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\zeta^2}\left(\lambda_t + \zeta_{i,t}\right) \text{ and } \frac{\sigma_\lambda^2}{\sigma_\zeta^2} = 2^{2\kappa} - 1.$$

The price level in period $t$ therefore equals

$$p_t = [(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]\left(1-2^{-2\kappa}\right)\lambda_t.$$

Hence, the set of rational expectations equilibria of the form $p_t = \theta\lambda_t$ for a given monetary policy $g_0 \in \mathbb{R}$ consists of the pairs $\left(\theta,\kappa^*\right) \in \mathbb{R}\times\mathbb{R}_+$ that solve the following two equations:

$$\theta = [(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]\left(1-2^{-2\kappa^*}\right), \tag{43}$$

and

$$\kappa^* = \arg\min_{\kappa\in\mathbb{R}_+} \frac{1}{1-\beta}\left\{\frac{\omega}{2}[(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]^2\sigma_\lambda^2 2^{-2\kappa} + f(\kappa)\right\}. \tag{44}$$

The first equation determines the response of the price level to markup shocks ($i.e.\ \theta$) as a function of the level of attention ($i.e.\ \kappa$). The second equation determines the equilibrium level of attention. Monetary policy enters in both equations.

Before deriving optimal monetary policy, one needs to make statements about existence and uniqueness of equilibrium for a given monetary policy $g_0 \in \mathbb{R}$. When the central bank completely offsets the effect of markup shocks on the profit-maximizing price (*i.e.* $g_0 = -\frac{\phi_\lambda}{\phi_c}$), there exists a unique solution to equations (43)–(44): The price level does not respond to markup shocks ($\theta = 0$) and price setters pay no attention to markup shocks ($\kappa = 0$). However, when the central bank does not fully offset the effect of markup shocks on the profit-maximizing price (*i.e.* $g_0 \neq -\frac{\phi_\lambda}{\phi_c}$), there may exist multiple solutions to equations (43)–(44). The reason is as follows. When price setters pay more attention to markup shocks, the price level responds more to markup shocks; and when the price level responds more to markup shocks and prices are strategic complements (*i.e.* $1-\phi_c > 0$), price setters have more incentive to pay attention to markup shocks. Whether this feedback effect is strong enough to generate multiplicity of equilibria depends on the degree of strategic complementarity in price setting and the degree of convexity of the cost function $f(\kappa)$. More precisely, the number of solutions to equations (43)–(44) depends on the monotonicity of the function $\varphi(\kappa)$ defined in the next proposition.[6]

**Proposition 3.** *(Existence and uniqueness of equilibrium for a given monetary policy) Let $\varphi(\kappa) = f'(\kappa)\left[\phi_c 2^\kappa + (1-\phi_c)2^{-\kappa}\right]^2$. If $\varphi$ is a strictly increasing function of $\kappa$ on $\mathbb{R}_+$, there exists a unique solution to equations (43)–(44) for each monetary policy $g_0 \in \mathbb{R}$. In contrast, if $\varphi$ is a non-increasing function of $\kappa$ on a subset of $\mathbb{R}_+$, there exist multiple solutions to equations (43)–(44) for some monetary policies $g_0 \in \mathbb{R}$.*

*Proof* See Appendix B. ||

Consider two examples. Suppose the cost of paying attention is linear in attention, $f(\kappa) = \mu\kappa$. Proposition 3 then implies that if $\phi_c \geq 1/2$ there exists a unique equilibrium for each monetary policy $g_0 \in \mathbb{R}$, while if $\phi_c < 1/2$ there exist multiple equilibria for some monetary policies $g_0 \in \mathbb{R}$. Next, suppose the cost of paying attention is proportional to $2^{2\kappa}$, $f(\kappa) = \mu 2^{2\kappa}$. In this case, Proposition 3 implies that there exists a unique equilibrium for each monetary policy $g_0 \in \mathbb{R}$, independent of the degree of strategic complementarity in price setting. The reason is that when the cost function for attention is more convex, a given increase in the benefit of paying attention leads to a smaller increase in the level of attention.

We can now turn to optimal monetary policy in the model with an endogenous information structure. Proposition 4 covers the case when there exists a unique equilibrium for each monetary policy $g_0 \in \mathbb{R}$ (*i.e.* $\varphi(\kappa)$ is strictly increasing on $\mathbb{R}_+$). Proposition 5 covers the case when there exist multiple equilibria for some policies $g_0 \in \mathbb{R}$ (*i.e.* $\varphi(\kappa)$ is non-increasing on a subset of $\mathbb{R}_+$). To shorten the proofs, we assume in both cases that $\varphi(\kappa)$ has no saddle point. One can then compute the derivative of equilibrium attention with respect to policy using the implicit function theorem.

**Proposition 4.** *(Endogenous information structure) Consider the Ramsey problem (21)–(37), where the signal precisions $1/\sigma_\eta^2$ and $1/\sigma_\zeta^2$ are given by the solution to problem (34)–(37). Assume $\sigma_a^2 = a_{-1} = 0$, $\sigma_\lambda^2 > 0$, and $\rho_\lambda = 0$. Consider policies of the form $G_t(L)\nu_t = g_0\nu_t$ and equilibria of the form $p_t = \theta\lambda_t$. Suppose that the function $\varphi(\kappa) = f'(\kappa)\left[\phi_c 2^\kappa + (1-\phi_c)2^{-\kappa}\right]^2$ is strictly increasing on $\mathbb{R}_+$ and has no saddle point on $\mathbb{R}_{++}$. Then, the following results hold. First, an*

6. The reason is that solving equation (43) for $\theta$ and substituting this equation into the first-order condition for the problem in equation (44) yields a condition of the form: constant equals $\varphi(\kappa)$, where the constant depends on $g_0$.

*optimal monetary policy has to satisfy*

$$g_0 \geq -\frac{\phi_\lambda}{\phi_c}. \tag{45}$$

*Second, take any policy $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$ at which price setters pay attention to markup shocks ($\kappa > 0$). At this policy, the derivative of price dispersion with respect to $g_0$ is positive and the derivative of equilibrium attention with respect to $g_0$ is positive. Furthermore, at this policy, the derivative of the consumption response to markup shocks $(g_0 - \theta)$ with respect to $g_0$ is non-positive if and only if*

$$\frac{f''(\kappa)}{f'(\kappa)} \leq 2\ln(2). \tag{46}$$

*Third, if the function $f(\kappa)$ satisfies condition (46) on $\mathbb{R}_{++}$, the unique optimal monetary policy is*

$$g_0^* = \begin{cases} -\frac{\phi_\lambda}{\phi_c} + \frac{\phi_\lambda}{\phi_c}\sqrt{\frac{f'(0)}{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}} & \text{if } \omega\phi_\lambda^2\sigma_\lambda^2\ln(2) > f'(0) \\ 0 & \text{otherwise} \end{cases}. \tag{47}$$

*At this policy, price setters pay no attention to markup shocks, the price level does not respond to markup shocks, price dispersion equals zero, and consumption variance equals $g_0^2\sigma_\lambda^2$.*

*Proof* See Appendix C. ‖

To understand the results in Proposition 4, note first what happens when the central bank does not respond with the monetary policy instrument to markup shocks (*i.e.* $g_0 = 0$). If at the inaction policy price setters pay no attention to markup shocks ($\omega\phi_\lambda^2\sigma_\lambda^2\ln(2) \leq f'(0)$), the inaction policy achieves the efficient allocation. Since price setters pay no attention to markup shocks, the price level does not respond to markup shocks and there is no inefficient price dispersion due to noise in the signal concerning the desired markup. Furthermore, since the policy instrument and the price level do not respond to markup shocks, consumption does not respond to markup shocks. The inaction policy achieves the first-best allocation. In addition, any other monetary policy yields inefficient consumption variance or inefficient price dispersion. Hence, if at the inaction policy price setters pay no attention to markup shocks ($\omega\phi_\lambda^2\sigma_\lambda^2\ln(2) \leq f'(0)$), the inaction policy is the unique optimal monetary policy.

If at the inaction policy price setters do pay attention to markup shocks ($\omega\phi_\lambda^2\sigma_\lambda^2\ln(2) > f'(0)$), the inaction policy does not achieve the efficient allocation. At the inaction policy, there is inefficient price dispersion (due to the noise in the signal concerning the desired markup) and inefficient consumption variance (because the price level strictly increases after a positive markup shock). To reduce inefficient price dispersion, the central bank can stabilize the profit-maximizing price in response to markup shocks with a policy $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, 0\right)$. Price setters then put a smaller weight on their noisy signals concerning the desired markup, which reduces inefficient price dispersion. This policy reduces the variance of the profit-maximizing price due to markup shocks. Price setters therefore pay less attention to markup shocks. Hence, the exact same policy that reduces inefficient price dispersion also makes price setters pay less attention to markup shocks! It is therefore unclear how this policy affects consumption variance. On the one hand, the contractionary monetary policy amplifies the fall in consumption after a positive markup shock. This effect is also present in the model with an exogenous information structure. On the other hand, the lower attention level reduces the response of the price level to markup shocks which reduces the fall in consumption after a positive markup shock. This effect is only present in the model with an endogenous information structure. If and only if the cost function for attention satisfies condition (46) at the current attention level, the second effect dominates at the current attention level. In this case, there is no trade-off between inefficient price dispersion and inefficient consumption variance: Decreasing one also decreases the other. Hence, if condition (46) holds for all positive attention levels, the optimal monetary policy is simple. Reduce the response of the profit-maximizing price to markup shocks until price setters just pay no attention to markup shocks. This policy is specified in the upper part of equation (47). At this policy, price setters pay no attention to markup shocks, the price level does not respond to markup shocks, price dispersion equals zero, and consumption variance equals $g_0^2\sigma_\lambda^2$. Going beyond this point by counteracting markup shocks even more strongly is suboptimal because price setters are already paying no attention to markup shocks and counteracting these shocks even more would only increase consumption variance.

In summary, if the cost of paying attention satisfies condition (46) for all attention levels, there exists no trade-off between inefficient price dispersion and inefficient consumption variance in the presence of markup shocks. Complete price stabilization in response to markup shocks is optimal.

Condition (46) has three remarkable features. First, the condition is very simple. It only involves the cost function for attention. Second, whether the trade-off between consumption variance and price dispersion disappears does not depend on the level of the cost of paying attention. The cost of paying attention can be arbitrarily small. Multiplying the cost function for attention $f(\kappa)$ with an arbitrarily small positive constant, say $10^{-100}$, does not affect whether condition (46) is satisfied. Third, the cost functions for attention that are most frequently used in the rational inattention literature, $f(\kappa)=\mu\kappa$ and $f(\kappa)=\mu 2^{2\kappa}$, satisfy condition (46) for all attention levels.

The next proposition covers the case when there exist multiple equilibria for some policies $g_0 \in \mathbb{R}$. Before one can make a statement about optimal monetary policy in this case, one has to make an assumption about the central bank's attitude towards multiple equilibria. Following the pertinent literature, we assume that the central bank dislikes policies that yield multiple equilibria. We consider two assumptions: (i) the central bank considers only policies $g_0 \in \mathbb{R}$ that yield a unique equilibrium, and (ii) the central bank considers policies that yield multiple equilibria but evaluates them by looking at the worst equilibrium. Both assumptions yield the same results concerning optimal monetary policy.

**Proposition 5.** *(Endogenous information structure) Consider the Ramsey problem (21)–(37), where the signal precisions $1/\sigma_\eta^2$ and $1/\sigma_\zeta^2$ are given by the solution to problem (34)–(37). Assume $\sigma_a^2 = a_{-1} = 0$, $\sigma_\lambda^2 > 0$, and $\rho_\lambda = 0$. Consider policies of the form $G_t(L)\nu_t = g_0\nu_t$ and equilibria of the form $p_t = \theta\lambda_t$. Suppose that the function $\varphi(\kappa) = f'(\kappa)\left[\phi_c 2^\kappa + (1-\phi_c)2^{-\kappa}\right]^2$ is non-increasing on a subset of $\mathbb{R}_+$ and has no saddle point on $\mathbb{R}_{++}$. Let $g_0^*$ denote the monetary policy given by equation (47) and let $\hat{g}_0$ denote the smallest $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \infty\right)$ at which there exist multiple equilibria. Then, the following results hold. First, an optimal monetary policy has to satisfy*

$$g_0 \geq -\frac{\phi_\lambda}{\phi_c}.$$

*Second, if the function $f(\kappa)$ satisfies condition (46) on $\mathbb{R}_{++}$, the optimal monetary policy is: $g_0 = g_0^*$ if $g_0^* < \hat{g}_0$, and a $g_0$ marginally below $\hat{g}_0$ if $\hat{g}_0 \leq g_0^*$. At this policy, price setters pay no attention to markup shocks, the price level does not respond to markup shocks, price dispersion equals zero, and consumption variance equals $g_0^2\sigma_\lambda^2$.*

*Proof* See Supplementary Appendix B. ‖

Moving from the unique equilibrium case to the multiple equilibria case changes hardly anything. The only change is that if at the policy $g_0^*$ given by equation (47) there exist multiple equilibria, then the optimal monetary policy is to counteract markup shocks even more strongly until no attention becomes the unique equilibrium.

We now study how welfare at the optimal monetary policy depends on the cost of paying attention. For ease of exposition, we focus on the unique equilibrium case. If the cost function for attention satisfies $f''(\kappa)/f'(\kappa) \leq 2\ln(2)$, consumption and price dispersion at the optimal policy equal $c_t = g_0^* \lambda_t$ and zero, respectively, where $g_0^*$ is given by equation (47). The value of the central bank's objective at the optimal policy equals

$$\sum_{t=0}^{\infty} \beta^t E\left[\left(c_t - c_t^*\right)^2 + \delta \frac{1}{I} \sum_{i=1}^{I} \left(p_{i,t} - p_t\right)^2\right] = \frac{1}{1-\beta}\left(g_0^*\right)^2 \sigma_\lambda^2.$$

Comparing the consumption response to markup shocks under endogenous information and optimal policy, $g_0^*$, to the consumption response to markup shocks under perfect information, $-\phi_\lambda/\phi_c$, shows that the equilibrium with endogenous information and optimal policy is better than the perfect information equilibrium, because the inefficient response of consumption to markup shocks is reduced or even suppressed entirely. Furthermore, reducing the cost of paying attention to aggregate conditions by giving price setters easier access to aggregate information reduces welfare. Formally, multiplying the cost function for attention $f(\kappa)$ with a constant $\tau \in [0,1)$ reduces welfare. The reason is that the central bank now has to counteract markup shocks even more to discourage price setters from paying attention to markup shocks. The more basic point is that having a flexible price level in response to markup shocks is bad not good. Finally, note that as $\tau \to 0$, the allocation under endogenous information and optimal policy converges to the perfect information allocation. Formally, as $\tau \to 0$, the consumption response to markup shocks converges to $-\phi_\lambda/\phi_c$ from above and price dispersion always equals zero at the optimal monetary policy.

### 5.2. *Autocorrelated desired markup*

In the case of an autocorrelated desired markup (*i.e.* $\rho_\lambda > 0$), we solve for the optimal monetary policy numerically.

Let us again begin with the model with an exogenous information structure. To solve the Ramsey problem (21)–(33) numerically, we turn this infinite-dimensional problem into a finite-dimensional problem by restricting $G_t(L)$ to be the same in each period and by restricting $G_t(L)$ to be the lag polynomial of an ARMA(2,2) process.[7] Following the procedure in Woodford (2002), one can then compute an exact linear rational expectations equilibrium of the model (22)–(33) *for a given monetary policy* by solving a Riccati equation. We then run a numerical optimization routine to obtain the optimal monetary policy.

We solved the Ramsey problem (21)–(33) for many different sets of parameter values. We always obtained the result that complete price stabilization in response to markup shocks is suboptimal. At the optimal monetary policy, there is inefficient price dispersion, and the price level strictly increases while consumption strictly falls on impact of a positive markup shock.[8]

7. We choose an ARMA(2,2) parameterization because it is well known from time series econometrics that an ARMA(p,q) parameterization is a very flexible and parsimonious parameterization.

8. For the optimal monetary policy at different parameter values, see Supplementary Appendix C.

Next let us turn to the model with an endogenous information structure. To solve the Ramsey problem (21)–(37) numerically, we again turn this infinite-dimensional problem into a finite-dimensional problem by restricting $G_t(L)$ to be a time-invariant lag polynomial of an ARMA(2,2) process. Following the procedure in Woodford (2002), one can then compute an exact linear rational expectations equilibrium of the model (22)–(33) for a given policy and a given signal precision. Furthermore, for a given law of motion for the endogenous variables, one can solve the attention problem (34)–(37). Hence, solving for a linear rational expectations equilibrium of the rational inattention model *for a given monetary policy* amounts to solving a fixed point problem. We solve for the optimal monetary policy both by evaluating the central bank's objective for different policies on a fine grid and by using an optimization routine.

We solved the Ramsey problem (21)–(37) for two different attention cost functions, $f(\kappa)=\mu\kappa$ and $f(\kappa)=\mu 2^{2\kappa}$, and many different sets of parameter values. We always obtained the result that, at the optimal monetary policy, price setters pay no attention to markup shocks, price dispersion equals zero, and the price level does not respond to markup shocks.[9]

## 6. ROBUSTNESS AND EXTENSIONS OF THE MAIN RESULTS

In this section, we cover the case of technology and markup shocks, and we go over many extensions of the model with an endogenous information structure (for a quick summary of the results, see Section 1).

### 6.1. *Technology and markup shocks*

This section covers the case when there are both technology and markup shocks. The result that complete price stabilization in response to markup shocks is optimal if condition (46) holds extends to this case.

When the cost function for attention is linear in attention, the Ramsey problem (21)–(37) separates into two independent problems: (i) finding the optimal monetary policy response to technology shocks, and (ii) finding the optimal monetary policy response to markup shocks. Propositions 1–5 then extend without any change to the case when there are both technology and markup shocks.

When the cost function for attention is strictly convex in attention, the two problems are connected through the cost function for attention: When the central bank increases the response of the profit-maximizing price to technology shocks, price setters pay more attention to technology shocks, which raises the marginal cost of paying attention to markup shocks. Let $\kappa_a$ and $\kappa_\lambda$ denote the attention allocated to technology shocks and markup shocks (*i.e.* they denote the first and second term on the right-hand side of equation (36)). For any given $\kappa_a \geq 0$, it is optimal for the central bank to make price setters pay no attention to markup shocks (*i.e.* $\kappa_\lambda = 0$). To see this, suppose $\kappa_\lambda > 0$. The central bank can then do the following. Lower the response of the profit-maximizing price to markup shocks, which reduces $\kappa_\lambda$. Simultaneously, lower the response of the profit-maximizing price to technology shocks so as to leave $\kappa_a$ unchanged if the fall in the marginal cost of paying attention due to the fall in $\kappa_\lambda$ would have led to an increase in $\kappa_a$. If the cost function for attention satisfies condition (46) for all $\kappa > \kappa_a$, the reduction in the response of the profit-maximizing price to markup shocks lowers inefficient price dispersion and inefficient consumption variance due to markup shocks. The proof is the same as before with the

9. In the benchmark model, price setters choose the precision of signals of the form "desired markup plus i.i.d. noise." We also solved for optimal monetary policy when price setters choose the precision of signals of the form "profit-maximizing price plus i.i.d. noise." The results stated above continue to hold. See Supplementary Appendix C.

only substantive change that $f'(0)$ is replaced by $f'(\kappa_a)$. At the same time, the reduction in the response of the profit-maximizing price to technology shocks reduces inefficient price dispersion and inefficient consumption variance due to technology shocks. Hence, if the cost function for attention satisfies condition (46) for all $\kappa > 0$, a policy that makes price setters pay attention to markup shocks cannot be optimal. For any given $\kappa_a \geq 0$, the optimal monetary policy makes price setters pay no attention to markup shocks.

The only remaining question is: What is the optimal level of $\kappa_a$? Put differently, what is the optimal response of the profit-maximizing price to technology shocks? On the one hand, reducing the response of the profit-maximizing price to technology shocks reduces inefficient price dispersion and inefficient consumption variance due to technology shocks. On the other hand, if $f''(\kappa) > 0$, a reduction in $\kappa_a$ lowers the marginal cost of paying attention to markup shocks. If and only if the second effect dominates, it is optimal to make price setters pay attention to technology shocks to raise the marginal cost of paying attention to markup shocks.

In summary, when there are both technology and markup shocks the following result still holds: If the cost function for attention satisfies $f''(\kappa)/f'(\kappa) \leq 2\ln(2)$ for all attention levels, there exists no trade-off between price dispersion and consumption variance in the presence of markup shocks and complete price stabilization in response to markup shocks is optimal. It is again optimal to make price setters pay no attention to markup shocks. The only change is that it may now be optimal to make price setters pay attention to technology shocks to raise the marginal cost of paying attention to markup shocks. Thus, complete price stabilization in response to technology shocks may be suboptimal. The difference to the model with an exogenous information structure becomes even starker. However, this new effect disappears once price setters can pay attention to any linear combination of $a_t$ and $\lambda_t$, which we turn to next.

### 6.2. *More general signal structure I: linear combinations*

In the benchmark model setup, it is assumed that paying attention to aggregate technology and paying attention to the desired markup are independent activities. We now assume that price setters can pay attention to any variable that is a linear combination of $a_t$ and $\lambda_t$.

The signal that the price setter in firm $i$ receives in period $t$ can be any signal of the form

$$s_{i,t} = \xi_a a_t + \xi_\lambda \lambda_t + \zeta_{i,t}. \tag{48}$$

The decision-maker chooses both the coefficients $(\xi_a, \xi_\lambda) \in \mathbb{R}^2$ and the signal precision $1/\sigma_\zeta^2 \in \mathbb{R}_+$.[10] The choice of coefficients can be interpreted as the choice of which variable to pay attention to, while the choice of signal precision can be interpreted as the choice of how much attention to devote to that variable.[11] In the new model with an endogenous information structure, the Ramsey problem (21)–(37) changes as follows: equation (48) replaces equation (28) and price setters choose $\xi_a$, $\xi_\lambda$, and $1/\sigma_\zeta^2$ instead of $1/\sigma_\eta^2$ and $1/\sigma_\zeta^2$. In addition, equation (17) replaces equation (36) because for the new signal structure equation (17) no longer simplifies to equation (36).

In the case of $\rho_a = \rho_\lambda = 0$, the optimal monetary policy is again the policy specified in Propositions 1, 4, and 5. The reason is quite simple. Consider first the attention problem of a price setter. When $\rho_a = \rho_\lambda = 0$, the price setter chooses to pay attention directly to the profit-maximizing price (*i.e.* it is optimal to choose the $(\xi_a, \xi_\lambda)$ with the property that $\xi_a a_t + \xi_\lambda \lambda_t$ equals

10. Adam (2007) and Mondria (2010) model the attention decision in a similar way.

11. When $\rho_a = \rho_\lambda = 0$ all endogenous variables are just linear functions of $a_t$ and $\lambda_t$ and thus equation (48) then implies that price setters can pay attention to any endogenous variable.

the equilibrium profit-maximizing price).[12] Furthermore, the optimal amount of attention devoted to the profit-maximizing price is the solution to

$$\min_{\kappa\in\mathbb{R}_+}\frac{1}{1-\beta}\left[\frac{\omega}{2}\sigma_{p^*}^2 2^{-2\kappa}+f(\kappa)\right], \tag{49}$$

where $\sigma_{p^*}^2$ denotes the variance of the profit-maximizing price. The signal-to-noise ratio of signal (48) equals $\frac{\sigma_{p^*}^2}{\sigma_\zeta^2}=2^{2\kappa}-1$ and the price of good $i$ equals $p_{i,t}=E\left[p_{i,t}^*|\mathcal{I}_{i,t}\right]=\left(1-2^{-2\kappa}\right)\left(p_{i,t}^*+\zeta_{i,t}\right)$.

Let us now turn to optimal monetary policy. First, consider the monetary policy response to aggregate technology shocks. Suppose that the central bank conducts the monetary policy specified in Proposition 1. This policy yields the efficient response of composite consumption to aggregate technology shocks and no response of the profit-maximizing price to aggregate technology shocks. Due to the second property, this policy is the monetary policy response to aggregate technology shocks that yields the smallest price dispersion and the smallest $\kappa$. Moreover, a small $\kappa$ is good because then prices also respond less to markup shocks. For these reasons, the optimal monetary policy response to aggregate technology shocks is the monetary policy specified in Proposition 1. Second, once the profit-maximizing price does not respond to aggregate technology shocks, problem (49) equals problem (42). Hence, the price setters' optimal allocation of attention is exactly the same as in the model with no aggregate technology shocks. Therefore, the optimal monetary policy response to markup shocks is the policy specified in Propositions 4 and 5.

### 6.3. *More general signal structure II: correlation in noise across firms*

In the benchmark model setup, it is assumed that all noise in signals is idiosyncratic. One can imagine that information processing noise is correlated across firms. Therefore, we now assume that the noise terms $\eta_{i,t}$ and $\zeta_{i,t}$ in equation (28) are correlated across firms with correlation coefficients $\rho_\eta\in[0,1]$ and $\rho_\zeta\in[0,1]$, respectively. Everything else remains unchanged. Let us begin with technology shocks. Proposition 1 extends to any $\rho_\eta\in[0,1]$. The reason is that the central bank can achieve the efficient allocation with $m_t=\frac{\phi_a}{\phi_c}a_t$ for any correlation in noise across firms.

Let us turn to markup shocks. We make the same assumptions as in Proposition 4 apart from three changes: (i) $\rho_\zeta\in[0,1]$, (ii) $p_t=\theta_\lambda\lambda_t+\theta_{\bar{\zeta}}\bar{\zeta}_t$ where $\bar{\zeta}_t$ is the common component of noise, and (iii) the function $\tilde{\varphi}(\kappa)=f'(\kappa)\left[\phi_c 2^\kappa+(1-\phi_c)\left(1-\rho_\zeta\right)2^{-\kappa}\right]^2$ is strictly increasing on $\mathbb{R}_+$ and has no saddle point on $\mathbb{R}_{++}$. Then, there exists at most one equilibrium with $\kappa>0$ at a given policy $g_0\in\mathbb{R}$. However, an equilibrium with $\kappa=0$ and an equilibrium with $\kappa>0$ can coexist. We make the same assumption about the central bank's attitude towards multiple equilibria as in Proposition 5. If the cost function for attention satisfies condition (46) on $\mathbb{R}_{++}$, the unique optimal monetary policy is

$$g_0^*=\begin{cases}-\frac{\phi_\lambda}{\phi_c}+\frac{\phi_\lambda}{\phi_c}\sqrt{\frac{f'(0)\left[1-(1-\phi_c)\rho_\zeta\right]}{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}} & \text{if } \omega\phi_\lambda^2\sigma_\lambda^2\ln(2)>f'(0)\left[1-(1-\phi_c)\rho_\zeta\right]\\ 0 & \text{otherwise}\end{cases}.$$

12. The argument is the same as the argument given in Maćkowiak and Wiederholt (2009), page 794.

At this policy price setters pay no attention to markup shocks. The proof is in Supplementary Appendix D.

### 6.4. *More general shocks*

The results for aggregate technology shocks extend to other shocks causing efficient fluctuations under perfect information. The results for markup shocks extend to other shocks causing inefficient fluctuations under perfect information.

To show this, we introduce a general exogenous aggregate variable: $z_t$. This variable affects the profit-maximizing price and may affect efficient composite consumption

$$p_{i,t}^{*} = p_t + \phi_c c_t + \phi_z z_t, \tag{50}$$

and

$$c_t^{*} = \vartheta z_t, \tag{51}$$

where $\phi_c > 0, \phi_z > 0$, and $\vartheta \in \mathbb{R}$. For ease of exposition, we assume that there is only one exogenous variable $z_t$ and the variable follows a Gaussian white noise process. In the new Ramsey problem, equations (50) and (51) replace equations (26) and (22), and $s_{i,t} = z_t + \zeta_{i,t}$ and $\kappa = \frac{1}{2}\log_2\left(\frac{\sigma_{z|t-1}^2}{\sigma_{z|t}^2}\right)$ replace equations (28) and (36). When price setters have perfect information, each firm sets the profit-maximizing price and thus $c_t = -\frac{\phi_z}{\phi_c} z_t$ and $p_{i,t} - p_t = 0$. The question is: What is the optimal monetary policy response to an innovation in $z_t$ when price setters have imperfect information?

In the model with an exogenous information structure, complete price stabilization in response to an innovation in $z_t$ is optimal if and only if the perfect-information response of the economy to an innovation in $z_t$ is efficient (*i.e.* $-\frac{\phi_z}{\phi_c} = \vartheta$). This is straightforward to show by following the steps in Appendix A.

Next consider the model with an endogenous information structure. We distinguish two cases: (i) $-\frac{\phi_z}{\phi_c} = \vartheta$, and (ii) $-\frac{\phi_z}{\phi_c} < \vartheta$. In the first case, the perfect-information response of the economy to an innovation in $z_t$ is efficient. In this case, the reasoning of Proposition 1 applies and thus the optimal monetary policy is $m_t = -\frac{\phi_z}{\phi_c} z_t$. In the second case, the perfect-information response of the economy to an innovation in $z_t$ is inefficient because the response is too large in magnitude ($-\frac{\phi_z}{\phi_c} < \vartheta \leq 0$) or has the wrong sign ($-\frac{\phi_z}{\phi_c} < 0 < \vartheta$). The proofs of Propositions 3 and 4 extend in a straightforward way from the desired markup to any variable $z_t$ with the property $-\frac{\phi_z}{\phi_c} < \vartheta$. Proposition 3 does not change at all. The only change in Proposition 4 apart from notation is that equation (47) generalizes to[13]

$$g_0^{*} = \begin{cases} -\frac{\phi_z}{\phi_c} + \frac{\phi_z}{\phi_c}\sqrt{\frac{f'(0)}{\omega\phi_z^2\sigma_z^2\ln(2)}} & \text{if } \omega(\phi_c\vartheta + \phi_z)^2\sigma_z^2\ln(2) > f'(0) \\ \vartheta & \text{otherwise} \end{cases}.$$

Hence, the results for markup shocks extend to any variable $z_t$ with the property that the perfect-information response of the economy to an innovation in $z_t$ is too large or has the wrong sign.

A simple example of a variable $z_t$ with the efficiency property is aggregate technology in the benchmark model setup ($z_t = -a_t$, $\phi_z = \phi_a$, and $\vartheta = -\frac{\phi_z}{\phi_c}$). A simple example of a variable $z_t$

13. The proof is in Supplementary Appendix E.

with the inefficiency property is the desired markup in that model ($z_t = \lambda_t$, $\phi_z = \phi_\lambda$, and $\vartheta = 0$). The results in this subsection are substantially more general. As an application of these results, consider what happens if we introduce real wage rigidity into the benchmark model setup. In the benchmark model setup, the real wage equals $w_t - p_t = \gamma c_t + \psi l_t$. Assume instead that the real wage equals $w_t - p_t = \varsigma[\gamma c_t + \psi l_t]$ with any $\varsigma \in [0,1)$.[14,15] Since the real wage responds less to consumption and labour supply, the coefficients on consumption and aggregate technology in equation (26) become

$$p^*_{i,t} = p_t + \frac{\varsigma\frac{\psi}{\alpha} + \varsigma\gamma + \frac{1-\alpha}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}} c_t - \frac{\varsigma\frac{\psi}{\alpha} + \frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}} a_t + \frac{\frac{\Lambda}{1+\Lambda}}{1 + \frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}} \lambda_t. \tag{52}$$

Otherwise the Ramsey problem remains unchanged; in particular, the efficient response of the economy to shocks does not change. Now the perfect-information response of composite consumption to technology shocks is too large in magnitude.[16] We immediately arrive at the following results. In the model with an exogenous information structure, complete price stabilization is *suboptimal* in response to technology *and* markup shocks. In contrast, in the model with an endogenous information structure, complete price stabilization is *optimal* in response to technology *and* markup shocks so long as condition (46) is satisfied. The reason is simple. Technology shocks now have the inefficiency property.

### 6.5. *Nominal wage rigidity*

In the previous section, we introduced real wage rigidity into the benchmark model setup. We now introduce nominal wage rigidity. We assume that the nominal wage rate equals its value in the non-stochastic steady state and households commit to supply any amount of labour at this wage rate. An information-based micro-foundation of this assumption is that households set the nominal wage rate one period in advance and the state of technology and the desired markup are i.i.d. over time. The difference to real wage rigidity is that the coefficient on the price level in equation (52) now differs from one, because movements in the price level cause movements in the real wage rate.

In the model with an exogenous information structure and nominal wage rigidity, complete price stabilization in response to technology shocks is never optimal. The reason is that the central bank uses movements in the price level to create movements in the real wage rate. Furthermore, complete price stabilization in response to markup shocks is never optimal.[17]

In contrast, in the model with an endogenous information structure and nominal wage rigidity, complete price stabilization is optimal both in response to technology shocks and in response to markup shocks so long as condition (46) holds and $f'(0)$ is large enough.[18] Complete price stabilization is again a more pervasive result in the model with an endogenous information structure.

14. Shimer (2012) proposes real wage rigidity as an explanation for the recent jobless recovery in the U.S.
15. When $\varsigma = 0$ one has to assume $\alpha < 1$ to ensure that the coefficient on consumption in equation (52) is non-zero.
16. Due to the real wage rigidity, households work too much in a boom and too little in a recession.
17. See Supplementary Appendix F for the proofs.
18. See Supplementary Appendix F. For parameter values that are standard in the business cycle literature and a marginal cost of attention in the range used in Maćkowiak and Wiederholt (2013) to match U.S. business cycle dynamics, complete price stabilization is optimal.

### 6.6. *Price setters who also take input decisions*

Angeletos and La'O (2012) show that complete price stabilization in response to technology shocks can be suboptimal when price setters have imperfect information and the same individuals who set prices also take input decisions. In their model the information structure is exogenous. We revisit this point in our model with an endogenous information structure.

Following Angeletos and La'O (2012), we now assume that the same individual who sets the price of good $i$ also takes an input decision. The representative household supplies two different types of labour, $L_{1,t}$ and $L_{2,t}$. The preferences of the representative household are given by

$$E_0\left[\sum_{t=0}^{\infty}\beta^t\left(\frac{C_t^{1-\gamma}-1}{1-\gamma}-\frac{L_{1,t}^{1+\psi_1}}{1+\psi_1}-\frac{L_{2,t}^{1+\psi_2}}{1+\psi_2}\right)\right],$$

where $\psi_1 \geq 0$ and $\psi_2 \geq 0$. The production function of firm $i$ is given by

$$Y_{i,t}=A_t L_{1,i,t}^{\alpha_1} L_{2,i,t}^{\alpha_2},$$

where $L_{1,i,t}$ and $L_{2,i,t}$ are firm $i$'s choice of the two labour inputs and $\alpha_1>0$ and $\alpha_2>0$ with $\alpha_1+\alpha_2\leq 1$. We consider two assumptions: (i) the individual who sets the price of good $i$ also chooses the level of one of the two labour inputs in every period, and (ii) the individual who sets the price of good $i$ chooses the labour mix $L_{1,i,t}/L_{2,i,t}$ in every period. Assumption (i) follows closely Angeletos and La'O (2012). The rest of the economy is modeled as in Section 2.

We first discuss optimal monetary policy in response to technology shocks. If the utility function is logarithmic in consumption ($\gamma=1$) or Frisch elasticities are identical ($\psi_1=\psi_2$) and price setters choose the labour mix, the optimal policy response to technology shocks is given by Proposition 1. Complete price stabilization in response to technology shocks is optimal, independent of whether the information structure is exogenous or endogenous. The reason is simple. If $\gamma=1$, the first-best labour inputs do not respond to technology shocks. If $\psi_1=\psi_2$, the first-best labour mix does not respond to technology shocks. The central bank can achieve the efficient allocation with $m_t=\frac{\phi_a}{\phi_c}a_t$. Moving away from these special cases, it matters whether information is exogenous or endogenous. If the information structure is exogenous and price setters choose the labour mix, complete price stabilization is *suboptimal* when $\gamma\neq 1$ and $\psi_1\neq\psi_2$. If the information structure is exogenous and price setters choose the level of one labour input, complete price stabilization is *suboptimal* when the constrained-efficient labour input responds to technology shocks, where constrained efficiency means that the planner cannot change the information structure. See Angeletos and La'O (2012). In contrast, in the model with an endogenous information structure, we find for the cost functions $f(\kappa)=\mu\kappa$ and $f(\kappa)=\mu 2^{2\kappa}$ that complete price stabilization in response to aggregate technology shocks is *optimal* so long as $\mu$ is large enough, independent of the values of the other parameters. Hence, in the model with endogenous information, complete price stabilization is once again a more pervasive result.[19]

In the model with an exogenous information structure, complete price stabilization in response to markup shocks is never optimal when price setters also take input decisions. In the model with an endogenous information structure, we find for the cost functions $f(\kappa)=\mu\kappa$ and $f(\kappa)=\mu 2^{2\kappa}$ that complete price stabilization in response to markup shocks is optimal so long as $\mu$ is large enough

19. See Supplementary Appendix G for details.

(or price setters choose the labour mix and $\psi_1 = \psi_2$). Hence, in the model with an endogenous information structure, complete price stabilization is again a more pervasive result.[20]

### 6.7. *Central bank information and interest rate rules*

This subsection presents results from varying assumptions about the central bank. First, so far we have assumed that the central bank directly controls nominal spending and commits to a nominal spending rule. In Supplementary Appendix H we show that the set of equilibria that the central bank can implement with an interest rate rule of the form (10) equals the set of equilibria that the central bank can implement with a nominal spending rule of the form (9). Second, so far we have assumed that the central bank has perfect information. In Supplementary Appendix H, we prove that when the central bank only receives a noisy signal about the desired markup $s_t = \lambda_t + \varrho_t$ (where the desired markup and the noise follow independent Gaussian white noise processes) and the central bank commits to a rule of the form $m_t = g_0 s_t$, the optimal monetary policy $g_0^*$ is still given by Proposition 4 if the information by the central bank is not too noisy (*i.e.* $\sigma_\varrho^2 > 0$ does not exceed a certain threshold).[21] At this policy, price setters pay no attention to markup shocks and therefore the price level does not respond to markup shocks. The only change is that welfare at the optimal monetary policy is weakly lower when the central bank has less precise information because in the case of $g_0^* \neq 0$ the noise in the central bank's signal introduces non-fundamental consumption variance.

## 7. CONCLUSION

This article solves a Ramsey optimal policy problem for an economy where decision-makers in firms choose how much attention they devote to aggregate conditions. When the allocation of attention by decision-makers in firms is exogenous, complete price stabilization is optimal *only* in response to shocks that cause *efficient* fluctuations under perfect information. When decision-makers in firms can choose how much attention they devote to aggregate conditions and the cost function for attention satisfies the simple condition $f''(\kappa)/f'(\kappa) \leq 2\ln(2)$, complete price stabilization is optimal *also* in response to shocks that cause *inefficient* fluctuations under perfect information. The optimality of complete price stabilization is a symptom of the fact that it is optimal to make price setters pay no attention to shocks that cause inefficient fluctuations. A corollary is that giving price setters easier access to information about the aggregate economy reduces welfare.

The result that it is optimal to make price setters pay no attention to markup shocks (and to other shocks causing inefficient fluctuations) is surprisingly robust. There are however exceptions to this result. First, the result requires that monetary policy has a sufficiently strong effect on the allocation of attention of price setters. This is the case if and only if $f''(\kappa)/f'(\kappa) \leq 2\ln(2)$. Second, if lack of attention by price setters to markup shocks causes other inefficiencies, then it can be suboptimal to make price setters pay no attention to markup shocks. This is the reason why complete price stabilization in response to markup shocks *can* be suboptimal when price setters also take input decisions (see Section 6.6). Third, if there is heterogeneity in the cost of paying attention across price setters and the mass of low-attention-cost price setters is sufficiently small,

20. For parameter values that are standard in the business cycle literature and for a value of the marginal cost of attention in the range used in Maćkowiak and Wiederholt (2013) to match U.S. business cycle dynamics, complete price stabilization is optimal in response to technology and markup shocks. See Supplementary Appendix G.

21. For a precise description of how to compute the threshold and the value of the threshold for different sets of parameter values, see Supplementary Appendix H.

it is optimal to make only the high-attention-cost price setters pay no attention to markup shocks. It is simply not worthwhile to make also the low-attention-cost price setters pay no attention to markup shocks due to their small mass.[22] The general point that survives even in the case of these exceptions is the following. Having a sticky price level in response to shocks that cause too large fluctuations under flexible prices is good not bad, and the central bank can affect price setters' incentive to pay attention to these shocks (i) by changing the response of the profit-maximizing price to these shocks and (ii) by changing the cost of paying attention to these shocks.

We think that interesting areas for future research are introducing rational inattention on the side of households and studying optimal fiscal policy in the case of endogenous attention.[23]

## APPENDIX

### A. PROOF OF PROPOSITION 2

**Step 1:** Substituting $c_t=m_t-p_t$, $m_t=g_0\lambda_t$, $p_t=\theta\lambda_t$, and $a_t=0$ into the equation for the profit-maximizing price (26) yields

$$p^*_{i,t}=[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]\lambda_t.$$

The price of good $i$ in period $t$ then equals

$$\begin{aligned} p_{i,t} &= [(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]E\left[\lambda_t|\mathcal{I}_{i,t}\right] \\ &= [(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]\frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\zeta^2}\left(\lambda_t+\zeta_{i,t}\right), \end{aligned} \tag{A.1}$$

and the price level in period $t$ equals

$$p_t=[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]\frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\zeta^2}\lambda_t. \tag{A.2}$$

Thus, the unique rational expectations equilibrium of the form $p_t=\theta\lambda_t$ is given by the solution to the following equation

$$\theta=[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]\frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\zeta^2}.$$

Solving the last equation for $\theta$ yields

$$\theta=\frac{\phi_c g_0+\phi_\lambda}{\phi_c+\frac{\sigma_\zeta^2}{\sigma_\lambda^2}}. \tag{A.3}$$

Substituting equation (A.3) into equations (A.1) and (A.2) yields

$$p_t = \frac{\phi_c g_0+\phi_\lambda}{\phi_c+\frac{\sigma_\zeta^2}{\sigma_\lambda^2}}\lambda_t, \tag{A.4}$$

$$p_{i,t}-p_t = \frac{\phi_c g_0+\phi_\lambda}{\phi_c+\frac{\sigma_\zeta^2}{\sigma_\lambda^2}}\zeta_{i,t}. \tag{A.5}$$

Substituting the monetary policy $m_t=g_0\lambda_t$ and equation (A.4) into $c_t=m_t-p_t$ yields

$$c_t=\frac{\frac{\sigma_\zeta^2}{\sigma_\lambda^2}g_0-\phi_\lambda}{\phi_c+\frac{\sigma_\zeta^2}{\sigma_\lambda^2}}\lambda_t. \tag{A.6}$$

The last three equations characterize the equilibrium for a given monetary policy $g_0\in\mathbb{R}$.

22. The extension with heterogeneity in the cost of paying attention is available from the authors upon request.

23. See Maćkowiak and Wiederholt (2013) for a dynamic stochastic general equilibrium (DSGE) model with rational inattention on the side of households and firms and a central bank following an exogenously given Taylor rule.

**Step 2:** Substituting equations (A.5), (A.6), (22) and $a_t = 0$ into the central bank's objective (21) yields

$$\frac{1}{1-\beta}\left[\left(\frac{\frac{\sigma_\zeta^2}{\sigma_\lambda^2}g_0-\phi_\lambda}{\phi_c+\frac{\sigma_\zeta^2}{\sigma_\lambda^2}}\right)^2\sigma_\lambda^2+\delta\left(\frac{\phi_c g_0+\phi_\lambda}{\phi_c+\frac{\sigma_\zeta^2}{\sigma_\lambda^2}}\right)^2\sigma_\zeta^2\right].$$

If $\sigma_\zeta^2 > 0$, the unique $g_0 \in \mathbb{R}$ that minimizes this expression is

$$g_0^* = \frac{(1-\delta\phi_c)\phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2}+\delta\phi_c^2}. \tag{A.7}$$

Substituting the optimal monetary policy $g_0^*$ into equations (A.4)–(A.6) yields

$$\begin{aligned} p_t &= \frac{\phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2}+\delta\phi_c^2}\lambda_t, \\ p_{i,t}-p_t &= \frac{\phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2}+\delta\phi_c^2}\zeta_{i,t}, \\ c_t &= -\frac{\delta\phi_c\phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2}+\delta\phi_c^2}\lambda_t. \end{aligned}$$

## B. PROOF OF PROPOSITION 3

First, solving equation (43) for $\theta$ yields

$$\theta = \frac{\left(1-2^{-2\kappa}\right)}{1-(1-\phi_c)\left(1-2^{-2\kappa}\right)}(\phi_c g_0+\phi_\lambda). \tag{B.8}$$

Second, the first-order condition for the attention problem in equation (44) reads

$$\omega[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]^2\sigma_\lambda^2\ln(2)2^{-2\kappa}\le f'(\kappa) \text{ with equality if } \kappa>0.$$

The left-hand side and the right-hand side of the weak inequality are the marginal benefit and the marginal cost of paying attention. Multiplying both sides of the weak inequality by $2^{2\kappa}$ yields

$$\omega[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]^2\sigma_\lambda^2\ln(2)\le 2^{2\kappa}f'(\kappa) \text{ with equality if } \kappa>0. \tag{B.9}$$

There exists a unique $\kappa \in \mathbb{R}_+$ satisfying this condition for a given $\theta \in \mathbb{R}$, because $2^{2\kappa}f'(\kappa)$ is a continuous and strictly increasing function on $\mathbb{R}_+$ and goes to infinity as $\kappa$ goes to infinity. Third, substituting equation (B.8) into equation (B.9) and rearranging yields

$$\omega(\phi_c g_0+\phi_\lambda)^2\sigma_\lambda^2\ln(2)\le f'(\kappa)\left[\phi_c 2^{\kappa}+(1-\phi_c)2^{-\kappa}\right]^2 \text{ with equality if } \kappa>0. \tag{B.10}$$

Thus, the pairs $(\theta,\kappa)\in\mathbb{R}\times\mathbb{R}_+$ solving equations (43)–(44) are given by the set of $\kappa\in\mathbb{R}_+$ satisfying condition (B.10) and the corresponding $\theta\in\mathbb{R}$ given by equation (B.8). The right-hand side of the weak inequality in condition (B.10) is a continuous function of $\kappa$ on $\mathbb{R}_+$ and goes to infinity as $\kappa$ goes to infinity. It follows from the Intermediate Value Theorem that there exists a $\kappa\in\mathbb{R}_+$ satisfying this condition. Moreover, if the right-hand side is a strictly increasing function of $\kappa$ on $\mathbb{R}_+$, then for each policy $g_0\in\mathbb{R}$ there exists a unique $\kappa\in\mathbb{R}_+$ satisfying this condition. In contrast, if the right-hand side is a non-increasing function of $\kappa$ on a subset of $\mathbb{R}_+$, then for some policies $g_0\in\mathbb{R}$ there exist multiple $\kappa\in\mathbb{R}_+$ satisfying this condition.

## C. PROOF OF PROPOSITION 4

**Step 1: Equilibrium for a given policy.** When the desired markup is i.i.d. over time ($\rho_\lambda = 0$), there are no technology shocks ($\sigma_a^2 = a_{-1} = 0$), and the central bank responds contemporaneously to markup shocks ($m_t = g_0\lambda_t$), the set of rational expectations equilibria of the form $p_t = \theta\lambda_t$ for a given monetary policy $g_0 \in \mathbb{R}$ consists of the pairs $(\theta,\kappa) \in \mathbb{R}\times\mathbb{R}_+$ solving equations (43)–(44). Solving equation (43) for $\theta$ yields

$$\theta = \frac{\left(1-2^{-2\kappa}\right)}{1-(1-\phi_c)\left(1-2^{-2\kappa}\right)}(\phi_c g_0 + \phi_\lambda). \tag{C.11}$$

Furthermore, the first-order condition for the attention problem in equation (44) reads

$$\omega[(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]^2 \sigma_\lambda^2 \ln(2) \le 2^{2\kappa} f'(\kappa) \text{ with equality if } \kappa > 0. \tag{C.12}$$

Substituting equation (C.11) into equation (C.12) and rearranging yields

$$\omega(\phi_c g_0 + \phi_\lambda)^2 \sigma_\lambda^2 \ln(2) \le f'(\kappa)\left[\phi_c 2^{\kappa} + (1-\phi_c)2^{-\kappa}\right]^2 \text{ with equality if } \kappa > 0. \tag{C.13}$$

Hence, the pairs $(\theta,\kappa) \in \mathbb{R}\times\mathbb{R}_+$ solving equations (43)–(44) are given by the set of $\kappa \in \mathbb{R}_+$ satisfying condition (C.13) and the corresponding $\theta \in \mathbb{R}$ given by equation (C.11). Next, it follows from equations (23), (25)–(28), and (36) that consumption, the profit-maximizing price of good $i$, and the actual price of good $i$ for a given $(g_0,\theta,\kappa) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}_+$ equal

$$c_t = (g_0 - \theta)\lambda_t$$

$$p^*_{i,t} = [(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]\lambda_t \tag{C.14}$$

$$p_{i,t} = [(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]\left(1-2^{-2\kappa}\right)\left(\lambda_t + \zeta_{i,t}\right), \tag{C.15}$$

where

$$\frac{\sigma_\lambda^2}{\sigma_\zeta^2} = 2^{2\kappa} - 1. \tag{C.16}$$

**Step 2: Optimal monetary policy has to satisfy** $g_0 \ge -\frac{\phi_\lambda}{\phi_c}$. At the monetary policy $g_0 = -\frac{\phi_\lambda}{\phi_c}$, the profit-maximizing price does not respond to markup shocks. The unique equilibrium is then $(\theta,\kappa) = (0,0)$, implying $E\left[\left(p_{i,t}-p_t\right)^2\right] = 0$ and $E\left[c_t^2\right] = \left(\frac{\phi_\lambda}{\phi_c}\right)^2 \sigma_\lambda^2$. Next, consider a monetary policy $g_0 < -\frac{\phi_\lambda}{\phi_c}$. Price dispersion at the policy $g_0 < -\frac{\phi_\lambda}{\phi_c}$ is weakly larger than price dispersion at the policy $g_0 = -\frac{\phi_\lambda}{\phi_c}$ because price dispersion is non-negative. Furthermore, consumption variance at the policy $g_0 < -\frac{\phi_\lambda}{\phi_c}$ is strictly larger than consumption variance at the policy $g_0 = -\frac{\phi_\lambda}{\phi_c}$ because for all $g_0 < -\frac{\phi_\lambda}{\phi_c}$ we have

$$g_0 - \frac{\left(1-2^{-2\kappa}\right)}{1-(1-\phi_c)\left(1-2^{-2\kappa}\right)}(\phi_c g_0 + \phi_\lambda) < -\frac{\phi_\lambda}{\phi_c} < 0.$$

Hence, a monetary policy $g_0 < -\frac{\phi_\lambda}{\phi_c}$ cannot be optimal. This means a monetary policy that makes the profit-maximizing price fall after a positive markup shock cannot be optimal. See equations (C.11) and (C.14). From now on, we focus on policies $g_0 \ge -\frac{\phi_\lambda}{\phi_c}$.

**Step 3: Equilibrium attention as a function of policy.** Let $\varphi(\kappa)$ denote the right-hand side of the weak inequality in condition (C.13), that is,

$$\varphi(\kappa) = f'(\kappa)\left[\phi_c 2^{\kappa} + (1-\phi_c)2^{-\kappa}\right]^2. \tag{C.17}$$

If $\varphi(\kappa)$ is a strictly increasing function on $\mathbb{R}_+$, there exists a unique equilibrium for each monetary policy $g_0 \in \mathbb{R}$. See Proposition 3. The unique equilibrium allocation of attention is

$$\kappa = \begin{cases} \kappa_{foc} & \text{if } \omega(\phi_c g_0 + \phi_\lambda)^2 \sigma_\lambda^2 \ln(2) > f'(0) \\ 0 & \text{otherwise} \end{cases}, \tag{C.18}$$

where $\kappa_{foc}$ is defined implicitly by

$$\omega(\phi_c g_0 + \phi_\lambda)^2 \sigma_\lambda^2 \ln(2) = \varphi\left(\kappa_{foc}\right). \tag{C.19}$$

If policy induces price setters to pay attention to markup shocks (*i.e.* $\omega(\phi_c g_0 + \phi_\lambda)^2 \sigma_\lambda^2 \ln(2) > f'(0)$ and thus $\kappa > 0$), then $\kappa = \kappa_{foc}$ and the implicit function theorem as well as $\varphi'(\kappa) \neq 0$ yields

$$\frac{\partial\kappa}{\partial g_0} = \frac{\omega 2(\phi_c g_0 + \phi_\lambda)\phi_c \sigma_\lambda^2 \ln(2)}{\varphi'(\kappa)}. \tag{C.20}$$

Hence, for any monetary policy $g_0 \ge -\frac{\phi_\lambda}{\phi_c}$ that induces price setters to pay attention to markup shocks, we have $\frac{\partial\kappa}{\partial g_0} > 0$.

**Step 4: Equilibrium price dispersion as a function of policy.** If price setters pay no attention to markup shocks ($\kappa = 0$), price dispersion equals zero. If price setters do pay attention to markup shocks ($\kappa > 0$), price dispersion is positive. More precisely, when price setters pay attention to markup shocks, it follows from equations (C.15)–(C.16) and (24) that the relative price of good $i$ equals

$$p_{i,t} - p_t = [(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]\left(1 - 2^{-2\kappa}\right)\zeta_{i,t},$$

and price dispersion equals

$$E\left[\left(p_{i,t} - p_t\right)^2\right] = [(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]^2\left(1 - 2^{-2\kappa}\right)^2 \frac{\sigma_\lambda^2}{2^{2\kappa} - 1}.$$

Furthermore, when price setters pay attention to markup shocks, it follows from equation (C.12) that equilibrium attention is given by

$$\omega[(1-\phi_c)\theta + \phi_c g_0 + \phi_\lambda]^2 \sigma_\lambda^2 \ln(2) = 2^{2\kappa} f'(\kappa).$$

Substituting the last equation into the previous equation yields

$$E\left[\left(p_{i,t} - p_t\right)^2\right] = \frac{f'(\kappa)\left(1 - 2^{-2\kappa}\right)}{\omega \ln(2)}.$$

Recall that for any monetary policy $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$ that induces price setters to pay attention to markup shocks, we have $\frac{\partial \kappa}{\partial g_0} > 0$. See Step 3. Hence, for any monetary policy $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$ that induces price setters to pay attention to markup shocks, the first derivative of equilibrium price dispersion with respect to $g_0$ is positive.

**Step 5: Equilibrium consumption as a function of policy.** Equilibrium consumption equals

$$c_t = (g_0 - \theta)\lambda_t.$$

Thus, how the response of consumption to markup shocks varies with policy depends on how the response of the price level to markup shocks varies with policy.

**Step 6: Equilibrium price level as a function of policy.** The price level equals $p_t = \theta\lambda_t$ with $\theta$ given by equation (C.11). If policy induces price setters to pay no attention to markup shocks ($\kappa = 0$), then $\theta = 0$. In contrast, if policy induces price setters to pay attention to markup shocks ($\kappa > 0$), then $\theta > 0$ and equation (C.11) implies

$$\frac{\partial \theta}{\partial g_0} = \frac{2^{-2\kappa} 2\ln(2)}{\left[1 - (1-\phi_c)\left(1 - 2^{-2\kappa}\right)\right]^2} \frac{\partial \kappa}{\partial g_0}(\phi_c g_0 + \phi_\lambda) + \frac{\left(1 - 2^{-2\kappa}\right)}{1 - (1-\phi_c)\left(1 - 2^{-2\kappa}\right)}\phi_c,$$

where $\frac{\partial \kappa}{\partial g_0}$ is given by equation (C.20). The first term is the attention effect that is only present in the model with an endogenous information structure. The second term is the usual effect which is also present in the model with an exogenous information structure. The second term is smaller than one and converges to one as $\kappa$ goes to infinity. Substituting equation (C.20) into the last equation yields

$$\frac{\partial \theta}{\partial g_0} = \frac{2^{-2\kappa} 2\ln(2)}{\left[1 - (1-\phi_c)\left(1 - 2^{-2\kappa}\right)\right]^2} \frac{\omega 2\phi_c \sigma_\lambda^2 \ln(2)}{\varphi'(\kappa)}(\phi_c g_0 + \phi_\lambda)^2 + \frac{\left(1 - 2^{-2\kappa}\right)}{1 - (1-\phi_c)\left(1 - 2^{-2\kappa}\right)}\phi_c.$$

Furthermore, substituting equations (C.17)–(C.19) into the last equation yields

$$\frac{\partial \theta}{\partial g_0} = \frac{4\ln(2)\phi_c f'(\kappa)}{\varphi'(\kappa)} + \frac{\left(1 - 2^{-2\kappa}\right)}{1 - (1-\phi_c)\left(1 - 2^{-2\kappa}\right)}\phi_c.$$

Rearranging the right-hand side of the last equation yields

$$\begin{aligned}
\frac{\partial \theta}{\partial g_0} &= \frac{4\ln(2)\phi_c f'(\kappa)}{\varphi'(\kappa)} - \frac{2^{-2\kappa}}{1 - (1-\phi_c)\left(1 - 2^{-2\kappa}\right)} + 1 \\
&= \frac{4\ln(2)\phi_c f'(\kappa)}{\varphi'(\kappa)} - \frac{1}{\phi_c 2^{2\kappa} + 1 - \phi_c} + 1 \\
&= \frac{4\ln(2)\phi_c f'(\kappa)\left[\phi_c 2^{2\kappa} + 1 - \phi_c\right] - \varphi'(\kappa)}{\varphi'(\kappa)\left[\phi_c 2^{2\kappa} + 1 - \phi_c\right]} + 1.
\end{aligned}$$

The definition of the function $\varphi(\kappa)$ implies

$$\varphi'(\kappa) = f''(\kappa)\left[\phi_c 2^\kappa + (1-\phi_c)2^{-\kappa}\right]^2 + f'(\kappa) 2\left[\phi_c^2 2^{2\kappa} - (1-\phi_c)^2 2^{-2\kappa}\right]\ln(2).$$

Substituting the last equation into the numerator of the previous equation and rearranging yields

$$\frac{\partial\theta}{\partial g_0}=\left[f'(\kappa)2\ln(2)-f''(\kappa)\right]\frac{\phi_c+(1-\phi_c)2^{-2\kappa}}{\varphi'(\kappa)}+1.$$

Hence, if policy induces price setters to pay attention to markup shocks ($\kappa>0$) and $\varphi'(\kappa)>0$, then $\frac{\partial\theta}{\partial g_0}\geq 1$ if and only if

$$\frac{f''(\kappa)}{f'(\kappa)}\leq 2\ln(2). \tag{C.21}$$

**Step 7: Optimal monetary policy.** Let us begin with the case $\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)\leq f'(0)$. In this case, at the policy $g_0=0$, price setters pay no attention to markup shocks, implying that price dispersion equals zero and consumption variance equals zero because $(g_0-\theta)^2\sigma_\lambda^2=\theta^2\sigma_\lambda^2=0$. Thus, the policy $g_0=0$ achieves the efficient allocation. In addition, any policy $g_0\neq 0$ does not achieve the efficient allocation. If at the policy $g_0\neq 0$ price setters pay no attention to markup shocks, consumption variance is positive because $(g_0-\theta)^2\sigma_\lambda^2=g_0^2\sigma_\lambda^2>0$. If at the policy $g_0\neq 0$ price setters do pay attention to markup shocks, price dispersion is positive. Hence, in the case $\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)\leq f'(0)$, the unique optimal monetary policy is $g_0=0$. At this policy, price setters pay no attention to markup shocks, the price level does not respond to markup shocks, and the equilibrium allocation equals the efficient allocation.

Let us turn to the case $\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)>f'(0)$. In this case, at the policy $g_0=0$, price setters do pay attention to markup shocks. Consider first equilibrium attention as a function of policy. Let $\bar{g}_0\geq-\frac{\phi_\lambda}{\phi_c}$ denote the policy at which price setters stop paying attention to markup shocks, that is, $\omega(\phi_c\bar{g}_0+\phi_\lambda)^2\sigma_\lambda^2\ln(2)=f'(0)$. The inequality $\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)>f'(0)$ implies that $\bar{g}_0<0$. For all $g_0\in\left[-\frac{\phi_\lambda}{\phi_c},\bar{g}_0\right]$ we have $\kappa=0$, while for all $g_0>\bar{g}_0$ we have $\kappa>0$. Now, consider price dispersion as a function of policy. For any policy $g_0\in\left[-\frac{\phi_\lambda}{\phi_c},\bar{g}_0\right]$ we have $\kappa=0$ and thus price dispersion equals zero, while for any policy $g_0>\bar{g}_0$ we have $\kappa>0$ and thus price dispersion is positive. Next, consider consumption variance as a function of policy. Note that $g_0-\theta$ is a continuous function of $g_0$ on $\left[-\frac{\phi_\lambda}{\phi_c},\infty\right)$, implying that consumption variance is a continuous function of $g_0$ on $\left[-\frac{\phi_\lambda}{\phi_c},\infty\right)$. Furthermore, consumption variance is a strictly decreasing function of $g_0$ on $\left[-\frac{\phi_\lambda}{\phi_c},\bar{g}_0\right]$, because for all $g_0\in\left[-\frac{\phi_\lambda}{\phi_c},\bar{g}_0\right]$ we have $\kappa=0$, $\theta=0$, $(g_0-\theta)^2\sigma_\lambda^2=g_0^2\sigma_\lambda^2$, and $g_0\leq\bar{g}_0<0$. Finally, if condition (C.21) holds for all $\kappa>0$, then consumption variance is a non-decreasing function of $g_0$ on $(\bar{g}_0,\infty)$, because at $g_0=\bar{g}_0$ we have $g_0-\theta=\bar{g}_0<0$, and for all $g_0>\bar{g}_0$ we have $\kappa>0$ and $\frac{\partial(g_0-\theta)}{\partial g_0}\leq 0$. See Step 6. Hence, if condition (C.21) holds for all $\kappa>0$, then in the case $\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)>f'(0)$ the unique optimal monetary policy is $g_0=\bar{g}_0$. This policy minimizes both price dispersion and consumption variance. At this policy, price setters pay no attention to markup shocks, the price level does not respond to markup shocks, price dispersion equals zero, and consumption variance equals $\bar{g}_0^2\sigma_\lambda^2$.

*Acknowledgments*. We thank Francesco Caselli, three anonymous referees, Fernando Alvarez, Marios Angeletos, Andy Atkeson, Larry Christiano, Jordi Gali, Christian Hellwig, Olivier Loisel, Bartosz Maćkowiak, Filip Matejka, Jordi Mondria, Jonathan Parker, Alessandro Pavan, Chris Sims, Aleh Tsyvinski, Pierre-Olivier Weill and seminar and conference participants at various institutions for helpful comments. An earlier version of this article was circulated under the title "Imperfect Information and Optimal Monetary Policy." Mirko Wiederholt thanks CREI and EIEF for hospitality.

**Supplementary Data**

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

ADAM, K. (2007), "Optimal Monetary Policy with Imperfect Common Knowledge", *Journal of Monetary Economics*, **54**, 267–301.

ANGELETOS, G.-M. and LA'O, J. (2012), "Optimal Monetary Policy with Informational Frictions" (Discussion Paper, MIT and University of Chicago).

BALL, L., MANKIW, G. N. and REIS, R. (2005), "Monetary Policy for Inattentive Economies", *Journal of Monetary Economics*, **52**, 703–725.

CALVO, G. (1983), "Staggered Prices in a Utility-Maximizing Framework", *Journal of Monetary Economics*, **12**, 383–398.

GALI, J. (2008), *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework*. (Princeton: Princeton University Press).

HELLWIG, C. and VELDKAMP, L. (2009), "Knowing What Others Know: Coordination Motives in Information Acquisition", *Review of Economic Studies*, **76**, 223–251.

LEEPER, E. (1991), "Equilibria under 'Active' and 'Passive' Monetary and Fiscal Policies", *Journal of Monetary Economics*, **27**, 129–147.

LORENZONI, G. (2010), "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information", *Review of Economic Studies*, **77**, 305–338.
LUO, Y. (2008), "Consumption Dynamics under Information Processing Constraints", *Review of Economic Dynamics*, **11**, 366–385.
MAĆKOWIAK, B. and WIEDERHOLT, M. (2009), "Optimal Sticky Prices under Rational Inattention", *American Economic Review*, **99**, 769–803.
MAĆKOWIAK, B. and WIEDERHOLT, M. (2013), "Business Cycle Dynamics under Rational Inattention" (Discussion Paper, ECB and Goethe University Frankfurt).
MATEJKA, F. (2011), "Rationally Inattentive Seller: Sales and Discrete Pricing" (Discussion Paper, CERGE-EI).
MONDRIA, J. (2010), "Portfolio Choice, Attention Allocation, and Price Comovement", *Journal of Economic Theory*, **145**, 1837–1864.
PACIELLO, L. (2012), "Monetary Policy and Price Responsiveness to Aggregate Shocks under Rational Inattention", *Journal of Money, Credit and Banking*, **44**, 1375–1399.
SHIMER, R. (2012), "Wage Rigidities and Jobless Recoveries", *Journal of Monetary Economics*, **59**, 65–77.
SIMS, C. A. (2003), "Implications of Rational Inattention", *Journal of Monetary Economics*, **50**, 665–690.
SIMS, C. A. (2006), "Rational Inattention: Beyond the Linear-Quadratic Case", *American Economic Review Papers and Proceedings*, **96**, 158–163.
SIMS, C. A. (2011), "Rational Inattention and Monetary Economics", in Friedman B.M. and Woodford, M. (eds.), *Handbook of Monetary Economics*, Vol. 3A (Amsterdam: Elsevier).
VAN NIEUWERBURGH, S. and VELDKAMP, L. (2009), "Information Immobility and the Home Bias Puzzle", *Journal of Finance*, **64**, 1187–1215.
VAN NIEUWERBURGH, S. and VELDKAMP, L. (2010), "Information Acquisition and Under-Diversification", *Review of Economic Studies*, **77**, 779–805.
WOODFORD, M. (2002), "Imperfect Common Knowledge and the Effects of Monetary Policy", in Aghion, P. *et al.* (eds.), *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, (Princeton and Oxford: Princeton University Press).
WOODFORD, M. (2003). *Interest and Prices. Foundations of a Theory of Monetary Policy.* (Princeton and Oxford: Princeton University Press).
WOODFORD, M. (2009). "Information-Constrained State-Dependent Pricing", *Journal of Monetary Economics*, **56**, 100–124.