Exogenous Information, Endogenous Information and Optimal Monetary Policy*

Luigi Paciello
Einaudi Institute for Economics and Finance

Mirko Wiederholt
Northwestern University

January 2011

Abstract

This paper 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.

JEL: E3, E5, D8.

Keywords: dispersed information, rational inattention, optimal monetary policy

*Paciello: Einaudi Institute for Economics and Finance, Via Sallustiana 62, 00187 Rome, Italy (e-mail: luigi.paciello@eief.it); Wiederholt: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (e-mail: m-wiederholt@northwestern.edu). We thank Fernando Alvarez, Andy Atkeson, Larry Christiano, Jordi Gali, Christian Hellwig, Olivier Loisel, Bartosz Maćkowiak, Jordi Mondria, Jonathan Parker, Alessandro Pavan, Aleh Tsyvinski, Pierre-Olivier Weill and seminar and conference participants at Bocconi, CREI, EIEF, HEC Montreal, IAE, Maryland, NAWMES 2011, NBER Summer Institute 2010, Northwestern, SED 2010, Toulouse and UCLA for helpful comments. An earlier version of this paper was circulated under the title "Imperfect Information and Optimal Monetary Policy." Mirko Wiederholt thanks CREI and EIEF for hospitality.

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 this question formally, we derive optimal monetary policy in two models. In the first model the amount of attention that decision-makers in firms allocate to aggregate conditions is exogenous. In the second model decision-makers in firms choose how much attention they devote to aggregate conditions. Our main findings concerning optimal monetary policy are the following. In the model with an exogenous allocation of attention by decision-makers in firms, complete price stabilization is the optimal policy *only* in response to shocks that cause *efficient* fluctuations under perfect information. In the model with an endogenous allocation of attention by decision-makers in firms, complete price stabilization is the optimal policy *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 monetary policy. The optimality of complete price stabilization becomes a much more general result.

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).[1] 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, we assume that slow adjustment of prices to changes in aggregate conditions is due to limited attention by decision-makers in firms rather than price stickiness à la Calvo (1983). Second, we endogenize the amount of attention that decision-makers in firms devote to aggregate conditions. We then address two questions: (1)

[1] For recent work on optimal monetary policy in New Keynesian models, see, e.g., Giannoni and Woodford (2010).

Does it matter for optimal monetary policy that slow adjustment of prices to changes in aggregate conditions is due to limited attention by decision-makers in firms rather than price stickiness à la Calvo (1983)? (2) Does it matter for optimal monetary policy whether the amount of attention that decision-makers in firms devote to aggregate conditions is exogenous or endogenous? The answer to the second question is the focus of this paper. This paper is the first paper solving a Ramsey optimal policy problem for an economy where decision-makers in firms choose how much attention they allocate to aggregate conditions.

We model the attention decision by decision-makers in firms following the literature on rational inattention (see Sims (2003)). Paying limited attention to aggregate conditions is modeled as receiving a noisy signal concerning aggregate conditions.[2] Paying more attention to aggregate conditions increases the precision of the signal. We assume that the noise in the signal is idiosyncratic because this accords well with the idea that the source of noise is limited attention by individual decision-makers rather than lack of publicly available information. Decision-makers in firms choose the amount of attention that they allocate to aggregate conditions facing an opportunity cost of allocating attention to aggregate conditions. The main prediction concerning the allocation of attention is that when the benefit of paying attention to aggregate conditions is larger, decision-makers in firms pay more attention to aggregate conditions.

In the model there are two types of shocks causing aggregate fluctuations: shocks that cause efficient fluctuations under perfect information and shocks that cause inefficient fluctuations under perfect information. In the benchmark model setup, these two types of shocks are aggregate technology shocks and markup shocks (i.e., shocks to the elasticity of substitution between goods). We start with aggregate technology shocks and markup shocks because studying the optimal monetary policy response to these shocks is common in the literature on optimal monetary policy. We then show that our results extend to other shocks. Furthermore, in the benchmark model setup, we assume that decision-makers in firms receive independent signals concerning aggregate technology and the desired markup. We also show that optimal monetary policy is identical when decision-makers in firms can decide to receive signals concerning any linear combination of aggregate technology and the desired markup (e.g., they can decide to receive signals/pay attention to endogenous variables).

[2] 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 United States was last year.

We derive optimal monetary policy under commitment assuming that the central bank aims to maximize expected utility of the representative household.

In the model with an exogenous information structure, complete price stabilization is optimal in response to aggregate technology shocks but not in response to markup shocks. The reason for the result about aggregate technology shocks is as follows. The response of the economy to aggregate technology shocks is efficient under perfect information. Furthermore, by stabilizing prices the central bank can replicate the perfect-information response of the economy to aggregate technology shocks. Thus, complete price stabilization is optimal in response to aggregate technology shocks. To understand the result concerning markup shocks, note first what happens when the monetary policy instrument (i.e., the money supply or nominal interest rate) remains constant after a markup shock. In this case, a positive markup shock (i.e., a shock that raises the desired markup) increases the profit-maximizing price. Price setters in firms therefore put a positive weight on their signals concerning the desired markup which causes inefficient price dispersion due to noise in the signal ("cross-sectional inefficiency"). Furthermore, the price level increases which given the constant monetary policy instrument causes a fall in consumption ("aggregate inefficiency"). To reduce cross-sectional inefficiency, the central bank can counteract the effect of the markup shock on the profit-maximizing price with a contractionary monetary policy (i.e., by lowering the money supply or raising the nominal interest rate). The profit-maximizing price then increases by less in response to a markup shock and price setters in firms therefore put less weight on their noisy signals concerning the desired markup, which reduces inefficient price dispersion. Unfortunately, the reduction in cross-sectional inefficiency comes at the cost of increased aggregate inefficiency: The contractionary monetary policy amplifies the fall in consumption. Hence, there is a trade-off between inefficient price dispersion and inefficient consumption variance. Furthermore, the marginal benefit of reducing inefficient price dispersion goes to zero as inefficient price dispersion goes to zero. Therefore, complete price stabilization in response to markup shocks is never optimal. This trade-off between cross-sectional inefficiency and aggregate inefficiency in the presence of markup shocks is emphasized a lot in the literature on optimal monetary policy and the result that complete price stabilization is not optimal in response to these shocks is a classic result in monetary economics.

In the model with an endogenous information structure, complete price stabilization is optimal in response to aggregate technology shocks *and* in response to markup shocks. This result is independent of parameter values. The reason for the result about markup shocks is the following. In the model with an endogenous information structure, decision-makers in firms pay more attention to aggregate conditions when the benefit of paying attention to aggregate conditions is larger. Consider again what happens when the central bank counteracts the effect of a positive markup shock on the profit-maximizing price with a contractionary monetary policy. There are two effects that are already present in the model with an exogenous information structure: The profit-maximizing price increases by less after a positive markup shock, implying that decision-makers in firms put less weight on their noisy signals, which reduces inefficient price dispersion; but the contractionary monetary policy by itself amplifies the fall in consumption. In addition, there is a new effect due to the endogenous allocation of attention. When the profit-maximizing price responds less to markup shocks, decision-makers in firms choose to pay less attention to the desired markup. Therefore, the price level increases by less after a positive markup shock, which by itself mutes the fall in consumption. It turns out that the new effect on consumption dominates for all parameter values. Thus, so long as decision-makers in firms pay some attention to the desired markup, the central bank can reduce both inefficient price dispersion and inefficient consumption variance by counteracting markup shocks more strongly. The classic trade-off between cross-sectional inefficiency and aggregate inefficiency disappears. The unique optimal monetary policy is the one that makes decision-makers in firms pay just no attention to the desired markup. Hence, at the optimal monetary policy, price setters in firms pay no attention to random variation in the desired markup and therefore prices do not respond to markup shocks. Complete price stabilization in response to markup shocks is optimal.

These results on optimal monetary policy in the model with an endogenous information structure generalize in important ways. Most importantly, the result that complete price stabilization is optimal in response to markup shocks extends to other shocks that cause inefficient fluctuations under perfect information and the result that complete price stabilization is optimal in response to aggregate technology shocks extends to other shocks that cause efficient fluctuations under perfect information. Aggregate technology shocks and markup shocks are just simple examples. Furthermore, these results on optimal monetary policy extend to more general signal structures. In the benchmark model setup, we assume that decision-makers in firms receive independent signals concerning aggregate technology and the desired markup and choose the precision of these two signals. In an extension, we assume that decision-makers in firms can choose to receive signals concerning any linear combination of aggregate technology and the desired markup (e.g., signals concerning endogenous variables). Optimal monetary policy is identical in these two model setups.

To summarize, let us answer the question that is the focus of this paper: Does it matter for optimal monetary policy whether the amount of attention that decision-makers in firms devote to aggregate conditions is exogenous or endogenous? Our answer is: a lot. When the amount of attention that decision-makers in firms devote to aggregate conditions is exogenous, complete price stabilization is the optimal policy *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 the optimal policy *also* in response to shocks that cause *inefficient* fluctuations under perfect information.

This paper is related to four recent papers studying optimal monetary policy in models with information frictions. The most closely related paper is Adam (2007). He studies optimal monetary policy in a model in which price setters in firms 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 labor supply shocks but not in response to markup shocks. Ball, Mankiw and Reis (2005) study optimal monetary policy in the sticky-information model of Mankiw and Reis (2002). In this model price setters in firms update their information 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. Finally, Lorenzoni (2010) and Angeletos and La'O (2008) study optimal monetary policy in models with dispersed information due to an island structure. In Lorenzoni (2010) price setters in firms observe the complete history of the economy up to the previous period, the sum of aggregate and idiosyncratic productivity, and a noisy public signal concerning aggregate productivity. There are several differences between his paper and our paper. In his paper the "noise" in the private signal concerning aggregate productivity is idiosyncratic productivity while in our paper the noise arises from limited attention; in his paper the information structure is exogenous whereas in our paper the information structure is endogenous; and in his paper the central bank has imperfect information. We initially assume that the central bank has perfect information about the state of the economy to derive the optimal monetary policy response to shocks, and we then study whether the central bank can also implement the optimal monetary policy with less information. Angeletos and La'O (2008) study optimal monetary policy when agents observe signals concerning endogenous variables with exogenous variance of noise. There is an information externality because a stronger response of agents to their private signals increases the signal-to-noise ratio of the signals concerning endogenous variables. Angeletos and La'O (2008) study how this information externality affects optimal fiscal and monetary policy. To recapitulate, the main difference between this paper and the papers cited above is that we derive optimal monetary policy when decision-makers in firms choose the amount of attention that they allocate to aggregate conditions.

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

The rest of the paper is organized as follows. Section 2 presents the model setup. Section 3 specifies the objective of the central bank. Section 4 states the optimal monetary policy problem under commitment in the model with an exogenous information structure and in the model with an endogenous information structure. Section 5 derives the equilibrium allocation under perfect information as a benchmark. Section 6 derives the optimal monetary policy response to aggregate technology shocks. Section 7 derives the optimal monetary policy response to markup shocks. Section 8 contains several additional results, including the results about more general shocks and more general signal structures. Section 9 concludes.

2 Model setup

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

Household: The household's preferences in period zero over sequences of consumption and labor 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 labor supply in period t. The parameter $\beta \in (0,1)$ is the discount factor, the parameter $\gamma > 0$ is the inverse of the intertemporal elasticity of substitution, and the parameter $\psi \geq 0$ is the inverse of the Frisch elasticity of labor supply. E_0 denotes the expectation operator conditioned on information of the household in period zero. Composite consumption in period t 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 consumption goods in period t equals $(1 + 1/\Lambda_t)$. We call the variable Λ_t the desired markup because Λ_t equals the desired markup by firms in period t. We assume that the log of the desired markup follows a stationary Gaussian first-order autoregressive process

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

where the parameter $\Lambda > 0$, the parameter $\rho_\lambda \in [0, 1)$, and the innovation ν_t is $i.i.d.N\left(0, \sigma_\nu^2\right)$. We call the innovation ν_t a markup shock. We introduce the markup shock in the model as an example of a shock that has the following property: The response of the economy to the shock under perfect information is inefficient.[3] In Section 7 we derive the optimal monetary policy response to markup shocks. In Section 8.4 we show that our results concerning markup shocks extend to other shocks that cause inefficient fluctuations under perfect information.

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

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

The right-hand side of the flow budget constraint is pre-consumption wealth in period t. Here B_{t-1} are the household's holdings of government bonds between period $t-1$ and period t, R_{t-1} is the nominal gross interest rate on those bond holdings, W_t is the nominal wage rate in period t, D_t are nominal aggregate profits in period t, T_t are nominal lump sum taxes in period t, and the term in brackets are unspent money balances carried over from period $t-1$ to period t. The representative household can transform pre-consumption wealth in period t into money balances, M_t, and bond holdings, B_t. The purpose of holding money is to purchase goods. We assume that the representative household faces the following cash-in-advance constraint

$$\sum_{i=1}^{I} P_{i,t} C_{i,t} = M_t. \tag{5}$$

[3] We define efficiency formally in Section 3.

The representative household also faces a no-Ponzi-scheme condition.

We introduce the cash-in-advance constraint because it allows us to explain the intuition for our results about optimal monetary policy in a simple way. In Section 8.6 we show that our results about optimal monetary policy extend to a cashless economy à la Woodford (2003). In addition, in Section 8.6 we study optimal monetary policy in a version of the economy with monetary transaction frictions. The formulation of the cash-in-advance constraint given above implies that there are no monetary transaction frictions because wage income can be transformed immediately into cash and cash can be spent immediately on goods. We decided to abstract from monetary transaction frictions in the benchmark economy for two reasons. First, abstracting from monetary transaction frictions is common in the New Keynesian literature on optimal monetary policy and thus abstracting from monetary transaction frictions facilitates comparison of our results to results about optimal monetary policy in the New Keynesian literature. Second, we think it is useful to study in isolation the implications of different frictions for optimal monetary policy. Therefore, we first abstract from monetary transaction frictions and we then add monetary transaction frictions in Section 8.6 by changing the timing of the cash-in-advance constraint, i.e., by assuming that cash has to be held for one period before it can be spent on goods.

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

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

$$Y_{i,t} = A_t L_{i,t}^{\alpha}, \tag{6}$$

where $Y_{i,t}$ is output and $L_{i,t}$ is labor 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 labor input. The log of aggregate productivity follows a stationary Gaussian first-order autoregressive process

$$\ln(A_t) = \rho_a \ln(A_{t-1}) + \varepsilon_t, \tag{7}$$

where the parameter $\rho_a \in [0,1)$ and the innovation ε_t is $i.i.d.N\left(0, \sigma_\varepsilon^2\right)$. We call the innovation ε_t an aggregate technology shock. The processes $\{A_t\}_{t=0}^{\infty}$ and $\{\Lambda_t\}_{t=0}^{\infty}$ are assumed to be independent. We introduce the aggregate technology shock in the model as an example of a shock that has the following property: The response of the economy to the shock under perfect information is efficient. In Section 6 we derive the optimal monetary policy response to aggregate technology shocks. In Section 8.4 we show that our results concerning aggregate technology shocks extend to other shocks that cause efficient fluctuations under perfect information.

Nominal profits of firm i in period t equal

$$\left(1+\tau_p\right) P_{i,t} Y_{i,t} - W_t L_{i,t}, \tag{8}$$

where τ_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 laws of motion for composite consumption, the nominal wage rate and the following price index[4]

$$P_t = \left(\frac{1}{I}\sum_{i=1}^{I} P_{i,t}^{-\frac{1}{\Lambda_t}}\right)^{-\Lambda_t} I. \tag{9}$$

Government: There is a monetary authority and a fiscal authority. The monetary authority commits to set the money supply according to the following rule

$$\ln\left(M_t^s\right) = F_t\left(L\right)\varepsilon_t + G_t\left(L\right)\nu_t, \tag{10}$$

where M_t^s denotes the money supply in period t. $F_t\left(L\right)$ and $G_t\left(L\right)$ are infinite-order lag polynomials which can depend on t. The last equation simply says that the log of the money supply in period t can be any linear function of the sequence of shocks up to and including period t. We will ask the question which linear function is optimal.

To study the optimal monetary policy response to shocks, we initially assume that the central bank has perfect information. In Section 8.7 we show that the central bank has a high incentive to be informed about the aggregate state of the economy and that the central bank can implement the optimal monetary policy also with less information. In addition, in Section 8.5 we show that the set of equilibria that the central bank can implement with a money supply rule of the form (10) equals the set of equilibria that the central bank can implement with an interest rate rule of

[4]Dixit and Stiglitz (1977), in their seminal article on monopolistic competition, also assume that there is a finite number of physical goods and that firms take the price index as given. Moreover, it seems to be a good description of the U.S. economy that there is a finite number of consumption goods and that firms take the consumer price index as given.

the form

$$\ln(R_t) = F_t(L)\,\varepsilon_t + G_t(L)\,\nu_t. \tag{11}$$

The drawback of committing to an interest rate rule rather than a money supply rule is that multiplicity of equilibria at a given monetary policy arises more easily. Therefore, we assume in the benchmark economy that the central bank can commit to a money supply rule and we postpone the discussion of unique implementation in the case of an interest rate rule to Section 8.5.[5]

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

$$T_t + B_t = R_{t-1}B_{t-1} + \tau_p\left(\sum_{i=1}^{I} P_{i,t}Y_{i,t}\right). \tag{12}$$

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. We assume that the fiscal authority pursues a Ricardian fiscal policy. In particular, for ease of exposition we assume that the fiscal authority fixes nominal government bonds at some non-negative level

$$B_t = B \geq 0. \tag{13}$$

Furthermore, we assume that the fiscal authority sets the production subsidy so as to correct the distortion arising from monopolistic competition in the non-stochastic steady state:

$$\tau_p = \Lambda. \tag{14}$$

Alternatively, one could assume that the fiscal authority sets the production subsidy so as to correct perfectly at each point in time the distortion arising from monopolistic competition:

$$\tau_{p,t} = \Lambda_t. \tag{15}$$

However, since in the United States fiscal policy has to be approved by Congress while monetary policy decisions are implemented directly by the Federal Reserve, we find it more realistic to assume that the fiscal authority *cannot* adjust the production subsidy quickly while the monetary authority *can* adjust the money supply quickly.

[5] A remark about the money market may be useful. In the model, the money market clears in the usual way. In equilibrium, endogenous variables (e.g., the price level, consumption and the nominal interest rate) adjust such that the demand for money balances by the representative household equals the supply of money balances by the monetary authority.

Information: We now specify the assumptions about the information structure. We consider two models. In the model with an exogenous information structure, the amount of attention that decision-makers in firms allocate to aggregate conditions is exogenous. In the model with an endogenous information structure, decision-makers in firms choose how much attention they allocate to aggregate conditions.

In both models, the information set of the price setter in firm i in period t is

$$\mathcal{I}_{i,t} = \mathcal{I}_{i,-1} \cup \{s_{i,0}, s_{i,1}, \ldots, s_{i,t}\}, \tag{16}$$

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. The latter is a two-dimensional vector consisting of a noisy signal concerning aggregate technology and a noisy signal concerning the desired markup:

$$s_{i,t} = \begin{pmatrix} \ln(A_t) + \eta_{i,t} \\ \ln(\Lambda_t/\Lambda) + \zeta_{i,t} \end{pmatrix}. \tag{17}$$

We assume that the noise in the signal is due to limited attention by the decision-maker.[6] The noise in the signal has the following properties: (i) the processes $\{\eta_{i,t}\}_{t=0}^{\infty}$ and $\{\zeta_{i,t}\}_{t=0}^{\infty}$ are independent of the processes $\{A_t\}_{t=0}^{\infty}$ and $\{\Lambda_t\}_{t=0}^{\infty}$, (ii) the processes $\{\eta_{i,t}\}_{t=0}^{\infty}$ and $\{\zeta_{i,t}\}_{t=0}^{\infty}$ 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 σ_η^2 and σ_ζ^2. The assumption that the noise in the signal is idiosyncratic accords well with the idea that the source of noise is limited attention by individual decision-makers rather than lack of publicly available information. The assumption that decision-makers in firms receive independent signals concerning aggregate technology and the desired markup is only for ease of exposition. In Section 8.3, we show that optimal monetary policy in the model with an endogenous information structure is identical when decision-makers in firms can decide to receive signals concerning any linear combination of aggregate technology and the desired markup (e.g., they can pay attention to endogenous variables).

In the model with an exogenous information structure, the variances of noise σ_η^2 and σ_ζ^2 are exogenous. In the model with an endogenous information structure, decision-makers in firms choose the precision of the signals facing an opportunity cost of allocating attention to aggregate conditions. Following the literature on rational inattention (see Sims (2003)), we quantify the amount

[6] See Footnote 2.

of attention allocated to aggregate conditions by uncertainty reduction. The timing is as follows. In period minus one, decision-makers in firms choose the precision of the signals so as to maximize expected profits net of the opportunity cost of devoting attention to aggregate conditions. In the following periods, decision-makers in firms receive the signals and take the optimal price setting decisions given the signals that they have received. Formally, the price setter in firm i solves the following decision problem in period minus one:

$$\max_{\left(1/\sigma_{\eta}^{2},1/\sigma_{\zeta}^{2}\right)\in\mathbb{R}_{+}^{2}}\left\{E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^{t}\pi\left(P_{i,t},P_{t},C_{t},W_{t},A_{t},\Lambda_{t}\right)\right]-\frac{\mu}{1-\beta}\kappa\right\}, \tag{18}$$

subject to equations (16)-(17) and

$$P_{i,t}=\arg\max_{P_{i}\in\mathbb{R}_{++}}E[\pi\left(P_{i},P_{t},C_{t},W_{t},A_{t},\Lambda_{t}\right)|\mathcal{I}_{i,t}], \tag{19}$$

and

$$\kappa=h\left(A_{t},\Lambda_{t}|\mathcal{I}_{i,t-1}\right)-h\left(A_{t},\Lambda_{t}|\mathcal{I}_{i,t}\right). \tag{20}$$

Consider first objective (18). 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 π denotes the real profit function defined as the nominal profit function divided by P_t times the marginal utility of consumption by the representative household. The variable κ is the amount of attention that the decision-maker devotes to aggregate conditions. The parameter $\mu > 0$ is the per-period marginal cost of devoting attention to aggregate conditions. We interpret the cost μ as an opportunity cost. Paying more attention to the price setting decision means paying less attention to some other activity. Following Sims (2003), we quantify the amount of attention devoted to aggregate conditions by uncertainty reduction, where uncertainty is measured by entropy. See equation (20). Here $h\left(A_t,\Lambda_t|\mathcal{I}_{i,t-1}\right)$ denotes the conditional entropy of A_t and Λ_t given $\mathcal{I}_{i,t-1}$ and $h\left(A_t,\Lambda_t|\mathcal{I}_{i,t}\right)$ denotes the conditional entropy of A_t and Λ_t given $\mathcal{I}_{i,t}$. The difference between the two quantifies the information received in period t. Finally, equation (19) specifies the price setting behavior. The basic trade-off is the following. A higher precision of the signals improves the price setting behavior but requires paying more attention to aggregate conditions.

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 the decision-maker has chosen the precision of the signals), the decision-maker receives 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 -1 equals the steady-state conditional covariance matrix of $(\ln(A_t), \ln(\Lambda_t))$ given information in period $t-1$.

Before we proceed, it is useful to point out two features of the decision problem (18)-(20) that one may expect to be important for our results about optimal monetary policy but that turn out to be irrelevant for our results about optimal monetary policy. First, in equation (17) we assume that decision-makers in firms receive independent signals concerning aggregate technology and the desired markup. In Section 8.3, we show that optimal monetary policy in the model with an endogenous information structure is identical when decision-makers in firms can decide to receive signals concerning any linear combination of aggregate technology and the desired markup (e.g., signals concerning endogenous variables). Second, in the decision problem (18)-(20) we assume that decision-makers in firms choose a constant signal precision once and for all. Our propositions about optimal monetary policy in the model with an endogenous information structure also hold when decision-makers in firms choose signal precision period by period or when decision-makers in firms choose signal precision as a function of time in period minus one.

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 about optimal monetary policy 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.

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 Objective of the central bank

We assume that the central bank aims to maximize expected utility of the representative household, given by equations (1)-(2).

We now derive a simple expression for expected utility of the representative household by using the fact that one can express period utility at a feasible allocation as a function only of the consumption vector at time t, aggregate productivity at time t, and the desired markup at time t. First, at any feasible allocation the representative household has to supply the labor that is needed to produce the consumption vector:

$$L_t = \sum_{i=1}^{I} \left(\frac{C_{i,t}}{A_t} \right)^{\frac{1}{\alpha}}. \tag{21}$$

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

$$1 = \frac{1}{I} \sum_{i=1}^{I} \hat{C}_{i,t}^{\frac{1}{1+\Lambda_t}},$$

where $\hat{C}_{i,t} = (C_{i,t}/C_t)$ is relative consumption of good i in period t. Rearranging yields

$$\hat{C}_{I,t} = \left(I - \sum_{i=1}^{I-1} \hat{C}_{i,t}^{\frac{1}{1+\Lambda_t}} \right)^{1+\Lambda_t}. \tag{22}$$

Substituting equations (21) and (22) into the period utility function in (1) yields the following expression for period utility at a feasible allocation

$$\begin{aligned} U\left(C_t, \hat{C}_{1,t}, \ldots, \hat{C}_{I-1,t}, A_t, \Lambda_t\right) &= \frac{C_t^{1-\gamma} - 1}{1-\gamma} \\ &\quad - \frac{1}{1+\psi} \left(\frac{C_t}{A_t} \right)^{\frac{1}{\alpha}(1+\psi)} \left[\sum_{i=1}^{I-1} \hat{C}_{i,t}^{\frac{1}{\alpha}} + \left(I - \sum_{i=1}^{I-1} \hat{C}_{i,t}^{\frac{1}{1+\Lambda_t}} \right)^{\frac{1}{\alpha}(1+\Lambda_t)} \right]^{1+\psi} \end{aligned} \tag{23}$$

Thus, expected utility at a feasible allocation equals

$$E\left[\sum_{t=0}^{\infty} \beta^t U\left(C_t, \hat{C}_{1,t}, \ldots, \hat{C}_{I-1,t}, A_t, \Lambda_t\right) \right].$$

To summarize, by substituting the technology and the consumption aggregator into the period utility function one can express period utility at a feasible allocation as a function only of the consumption vector at time t, aggregate productivity at time t, and the desired markup at time t.

We define the efficient allocation in period t as the feasible allocation in period t that maximizes utility of the representative household. The efficient allocation in period t is given by

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

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 period t and is independent of the desired markup. The efficient consumption mix in period t is to consume an equal amount of each good.

In the following sections, we work with a log-quadratic approximation to the period utility function (23) around the non-stochastic steady state. In the following, variables without time subscript denote values in the non-stochastic steady state and small variables denote log-deviations from the non-stochastic steady state (e.g., $c_t = \ln(C_t/C)$ and $\hat{c}_{i,t} = \ln\left(\hat{C}_{i,t}/\hat{C}_i\right)$). Due to the production subsidy (14), the non-stochastic steady state is efficient (i.e., $C = C^*$ and $\hat{C}_i = \hat{C}_i^*$). Expressing the period utility function U defined by equation (23) in terms of log-deviations from the non-stochastic steady state and using $C = C^*$ and $\hat{C}_i = \hat{C}_i^*$ yields the following expression for period utility at a feasible allocation

$$\begin{aligned}
& u\left(c_t, \hat{c}_{1,t}, \ldots, \hat{c}_{I-1,t}, a_t, \lambda_t\right) \\
= \quad & \frac{C^{1-\gamma}e^{(1-\gamma)c_t} - 1}{1-\gamma} \\
& -\frac{C^{1-\gamma}e^{\frac{1}{\alpha}(1+\psi)(c_t-a_t)}}{\frac{1}{\alpha}(1+\psi)} \left[\frac{1}{I}\sum_{i=1}^{I-1} e^{\frac{1}{\alpha}\hat{c}_{i,t}} + \frac{1}{I}\left(I - \sum_{i=1}^{I-1} e^{\hat{c}_{i,t}\frac{1}{1+\Lambda e^{\lambda_t}}}\right)^{\frac{1}{\alpha}\left(1+\Lambda e^{\lambda_t}\right)}\right]^{1+\psi}. \qquad (24)
\end{aligned}$$

A second-order Taylor approximation to this function at the non-stochastic steady state yields the result stated in Proposition 1.

Proposition 1 *(Objective of the central bank) Let $\tilde{u}$ denote the second-order Taylor approximation to the period utility function u at the origin. Let E denote the unconditional expectation operator.*

Let x_t, z_t, and ω_t denote the following vectors

$$\begin{aligned} x_t &= \begin{pmatrix} c_t & \hat{c}_{1,t} & \cdots & \hat{c}_{I-1,t} \end{pmatrix}', \\ z_t &= \begin{pmatrix} a_t & \lambda_t \end{pmatrix}', \\ \omega_t &= \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}'. \end{aligned}$$

Let $\omega_{n,t}$ denote the nth element of ω_t. Suppose that there exist two constants $\delta < (1/\beta)$ and $\phi \in \mathbb{R}$ such that, for each period $t \geq 0$ and for all n and k,

$$E\left|\omega_{n,t}\omega_{k,t}\right| < \delta^t \phi. \tag{25}$$

Then

$$E\left[\sum_{t=0}^{\infty}\beta^t \tilde{u}\left(x_t, z_t\right)\right] - E\left[\sum_{t=0}^{\infty}\beta^t \tilde{u}\left(x_t^*, z_t\right)\right] = \sum_{t=0}^{\infty}\beta^t E\left[\frac{1}{2}\left(x_t - x_t^*\right)' H \left(x_t - x_t^*\right)\right], \tag{26}$$

where the matrix H is given by

$$H = -C^{1-\gamma}\begin{bmatrix} \gamma - 1 + \frac{1}{\alpha}(1+\psi) & 0 & \cdots & \cdots & 0 \\ 0 & 2\frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \cdots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} \\ \vdots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} \\ 0 & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \cdots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & 2\frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} \end{bmatrix}, \tag{27}$$

and the vector x_t^* is given by

$$c_t^* = \frac{\frac{1}{\alpha}(1+\psi)}{\gamma - 1 + \frac{1}{\alpha}(1+\psi)} a_t, \tag{28}$$

and

$$\hat{c}_{i,t}^* = 0. \tag{29}$$

Proof. See Appendix A. ■

After the log-quadratic approximation to the period utility function (23) around the non-stochastic steady state, the efficient consumption vector in period t is given by equations (28)-(29) and the loss in expected utility in the case of deviations of the actual consumption vector from the efficient consumption vector is given by equation (26). The upper-left element of the matrix H determines the loss in utility in the case of an inefficient consumption level. The lower-right block of the matrix H determines the loss in utility in the case of an inefficient consumption mix. Condition (25) ensures that in the expressions on the left-hand side of equation (26) one can change the order of integration and summation and the infinite sum converges. In the models that we consider, condition (25) is always satisfied.

4 The Ramsey problem

In this section, we state the maximization problem of the central bank that aims to commit to the policy rule that maximizes expected utility of the representative household.

In the model with an exogenous information structure, the problem of the central bank is

$$\max_{\{F_t(L),G_t(L)\}_{t=0}^{\infty}} E\left[\sum_{t=0}^{\infty}\beta^t U\left(C_t,\hat{C}_{1,t},\ldots,\hat{C}_{I-1,t},A_t,\Lambda_t\right)\right], \tag{30}$$

subject to

$$P_tC_t = M_t, \tag{31}$$

$$C_{i,t} = \left(\frac{P_{i,t}}{\frac{1}{I}P_t}\right)^{-\left(1+\frac{1}{\Lambda_t}\right)} C_t, \tag{32}$$

$$\frac{W_t}{P_t} = L_t^{\psi} C_t^{\gamma}, \tag{33}$$

$$P_t = \left(\frac{1}{I}\sum_{i=1}^{I} P_{i,t}^{-\frac{1}{\Lambda_t}}\right)^{-\Lambda_t} I, \tag{34}$$

$$P_{i,t} = \arg\max_{P_i\in\mathbb{R}_{++}} E[\pi\left(P_i,P_t,C_t,W_t,A_t,\Lambda_t\right)|\mathcal{I}_{i,t}], \tag{35}$$

$$\mathcal{I}_{i,t} = \mathcal{I}_{i,-1}\cup\left\{s_{i,0},s_{i,1},\ldots,s_{i,t}\right\}, \tag{36}$$

$$s_{i,t} = \begin{pmatrix} \ln\left(A_t\right)+\eta_{i,t} \\ \ln\left(\Lambda_t/\Lambda\right)+\zeta_{i,t}\end{pmatrix}, \tag{37}$$

$$L_t = \sum_{i=1}^{I}\left(\frac{C_{i,t}}{A_t}\right)^{\frac{1}{\alpha}}, \tag{38}$$

$$\ln\left(A_t\right) = \rho_a \ln\left(A_{t-1}\right)+\varepsilon_t, \tag{39}$$

$$\ln\left(\Lambda_t/\Lambda\right) = \rho_\lambda \ln\left(\Lambda_{t-1}/\Lambda\right)+\nu_t, \tag{40}$$

and

$$\ln\left(M_t\right) = F_t\left(L\right)\varepsilon_t + G_t\left(L\right)\nu_t. \tag{41}$$

The objective (30) is expected utility of the representative household. The function U defined by equation (23) gives period utility at a feasible allocation. Equations (31)-(34) are the household's optimality conditions.[7] Equation (35) states that price setters in firms take the optimal price setting decisions given their information. Equations (36)-(37) specify the information set of the price setter in firm i in period t. Equation (38) is the labor market clearing condition. Equations (39)-(40) specify the laws of motion for the exogenous variables. Finally, equation (41) is the equation for the money supply, where $F_t(L)$ and $G_t(L)$ are infinite-order lag polynomials which can depend on t.[8] The innovations ε_t, ν_t, $\eta_{i,t}$ and $\zeta_{i,t}$ have the properties specified in Section 2. In the model with an exogenous information structure, the variances of noise σ_η^2 and σ_ζ^2 are structural parameters. They do not depend on monetary policy.

In the model with an endogenous information structure, the variances of noise σ_η^2 and σ_ζ^2 are given by the solution to the attention problem (18)-(20) and the central bank understands that the choice of the policy rule affects the firms' allocation of attention.

In the literature on optimal monetary policy it is common practice to study the Ramsey problem after a log-quadratic approximation of the central bank's objective and a log-linear approximation of the equilibrium conditions. See Woodford (2003), Gali (2008), Adam (2007) and Ball, Mankiw and Reis (2005). We follow this common practice. This makes our results comparable to their results. A 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 linear quadratic Ramsey problem in the model with an exogenous information structure:

$$\min_{\{F_t(L),G_t(L)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t E\left[(c_t - c_t^*)^2 + \delta \frac{1}{I}\sum_{i=1}^{I}(p_{i,t} - p_t)^2\right], \tag{42}$$

subject to

$$c_t^* = \frac{\phi_a}{\phi_c} a_t, \tag{43}$$

$$c_t = m_t - p_t, \tag{44}$$

$$p_t = \frac{1}{I}\sum_{i=1}^{I} p_{i,t}, \tag{45}$$

[7] We do not state the consumption Euler equation because here the consumption Euler equation is only a pricing equation determining the equilibrium nominal interest rate.

[8] The requirement that each firm produces the quantity demanded is embedded in the profit function π and the money market clearing condition $M_t = M_t^s$ is embedded in equation (41). See Footnote 5.

$$p_{i,t} = E\left[p^*_{i,t}|\mathcal{I}_{i,t}\right], \tag{46}$$

$$p^*_{i,t} = p_t + \phi_c c_t - \phi_a a_t + \phi_\lambda \lambda_t, \tag{47}$$

$$\mathcal{I}_{i,t} = \mathcal{I}_{i,-1} \cup \{s_{i,0}, s_{i,1}, \ldots, s_{i,t}\}, \tag{48}$$

$$s_{i,t} = \begin{pmatrix} a_t + \eta_{i,t} \\ \lambda_t + \zeta_{i,t} \end{pmatrix}, \tag{49}$$

$$a_t = \rho_a a_{t-1} + \varepsilon_t, \tag{50}$$

$$\lambda_t = \rho_\lambda \lambda_{t-1} + \nu_t, \tag{51}$$

and

$$m_t = F_t(L)\,\varepsilon_t + G_t(L)\,\nu_t, \tag{52}$$

where

$$\phi_c = \frac{\frac{\psi}{\alpha} + \gamma + \frac{1-\alpha}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}} > 0, \tag{53}$$

$$\phi_a = \frac{\frac{\psi}{\alpha} + \frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}} > 0, \tag{54}$$

$$\phi_\lambda = \frac{\frac{\Lambda}{1+\Lambda}}{1 + \frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}} > 0, \tag{55}$$

$$\delta = \frac{\frac{1+\Lambda-\alpha}{(1+\Lambda)\alpha}\left(1 + \frac{1}{\Lambda}\right)^2}{\gamma - 1 + \frac{1}{\alpha}(1+\psi)} > 0. \tag{56}$$

The objective (42) follows from substituting the log-linear demand function for good i into equation (26) and by using equation (45). The variable c^*_t is efficient composite consumption in period t and the parameter δ is the relative weight on cross-sectional inefficiency versus aggregate inefficiency in the central bank's objective. The variable $p^*_{i,t}$ is the profit-maximizing price of good i in period t.

In the model with an exogenous information structure, the variances of noise σ^2_η and σ^2_ζ are exogenous. In the model with an endogenous information structure, price setters in firms choose the precision of the signals. After a log-quadratic approximation of the real profit function π around the non-stochastic steady state, the attention problem (18)-(20) reads:

$$\min_{(1/\sigma^2_\eta, 1/\sigma^2_\zeta)\in\mathbb{R}^2_+} \left\{ E_{i,-1}\left[\sum_{t=0}^{\infty} \beta^t \frac{\omega}{2}\left(p_{i,t} - p^*_{i,t}\right)^2\right] + \frac{\mu}{1-\beta}\kappa \right\}, \tag{57}$$

subject to

$$p_{i,t} = E\left[p^*_{i,t}|\mathcal{I}_{i,t}\right], \tag{58}$$

and

$$\kappa = \frac{1}{2}\log_2\left(\frac{\sigma^2_{a|t-1}}{\sigma^2_{a|t}}\right) + \frac{1}{2}\log_2\left(\frac{\sigma^2_{\lambda|t-1}}{\sigma^2_{\lambda|t}}\right), \tag{59}$$

where the coefficient ω determining the profit loss in the case of a deviation of the actual price from the profit-maximizing price is given by

$$\omega = C^{-\gamma}\frac{WL_i}{P}\frac{\frac{1+\Lambda}{\Lambda}}{\alpha}\left(1 + \frac{1-\alpha}{\alpha}\frac{1+\Lambda}{\Lambda}\right). \tag{60}$$

Here $p^*_{i,t}$ denotes the profit-maximizing price of good i in period t given by equation (47) and $\mathcal{I}_{i,t}$ denotes the information set of the price setter in firm i in period t given by equations (48)-(49). Equation (20) reduces to equation (59) because the conditional distribution of (a_t, λ_t) is Gaussian both given $\mathcal{I}_{i,t-1}$ and given $\mathcal{I}_{i,t}$ and because a_t and λ_t are conditionally independent both given $\mathcal{I}_{i,t-1}$ and given $\mathcal{I}_{i,t}$. In the following, σ^2_λ, $\sigma^2_{\lambda|t-1}$ and $\sigma^2_{\lambda|t}$ denote the unconditional variance of λ_t, the conditional variance of λ_t given $\mathcal{I}_{i,t-1}$ and the conditional variance of λ_t given $\mathcal{I}_{i,t}$, respectively.

5 Perfect information solution

As a benchmark, we now derive the response of the economy to aggregate shocks under perfect information.

Suppose that price setters in firms have perfect information. Each firm then charges the profit-maximizing price and equations (44)-(47) imply

$$c_t = \frac{\phi_a}{\phi_c}a_t - \frac{\phi_\lambda}{\phi_c}\lambda_t, \tag{61}$$

$$p_{i,t} - p_t = 0, \tag{62}$$

and

$$p_t = m_t - c_t. \tag{63}$$

Under perfect information, the response of the economy to aggregate technology shocks is efficient while the response of the economy to markup shocks is inefficient. To see this, note that there is no inefficient price dispersion under perfect information and compare equations (43) and (61). 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.

6 Optimal monetary policy response to technology shocks

In this section, we derive the optimal monetary policy response to aggregate technology shocks in the model with an exogenous information structure and in the model with an endogenous information structure. We show that in both models complete price stabilization is the optimal policy in response to aggregate technology shocks.

6.1 Exogenous information structure

Proposition 2 *(Exogenous signal precision) Consider the Ramsey problem (42)-(56), where the variances of noise σ_η^2 and σ_ζ^2 are exogenous. Consider equilibria with the property that the price level p_t is a linear function of the shocks. If $\sigma_\eta^2 > 0$, the unique optimal monetary policy response to aggregate technology shocks is*

$$F_t(L)\,\varepsilon_t = \frac{\phi_a}{\phi_c} a_t. \tag{64}$$

At the optimal monetary policy, the price level does not respond to aggregate technology shocks.

Proof. See Appendix B. ■

The reason for this result about optimal policy is the following. The response of the economy to aggregate technology shocks under perfect information is efficient. Furthermore, by offsetting the effect of aggregate technology shocks on the profit-maximizing price the central bank can replicate the perfect-information response of real variables to aggregate technology shocks. To see this, note that the profit-maximizing price (47) can be written as

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

By setting $F_t(L)\,\varepsilon_t = \frac{\phi_a}{\phi_c} a_t$, the central bank can offset the effect of aggregate technology shocks on the profit-maximizing price. Price setters in firms then put no weight on their noisy signals concerning aggregate technology and thus there is no inefficient price dispersion due to the noise in the signal concerning aggregate technology. In addition, the price level then does not respond to aggregate technology shocks and therefore the response of the consumption level to aggregate technology shocks equals $\frac{\phi_a}{\phi_c} a_t$, which equals the efficient response of the consumption level to aggregate technology shocks. See equations (44) and (43).

6.2 Endogenous information structure

Proposition 3 *(Endogenous signal precision) Consider the Ramsey problem (42)-(60), where the signal precisions* $\left(1/\sigma_\eta^2\right)$ *and* $\left(1/\sigma_\zeta^2\right)$ *are given by the solution to problem (57)-(60). Consider equilibria with the property that the price level* p_t *is a linear function of the shocks. If* $\mu > 0$*, the unique optimal monetary policy response to aggregate technology shocks is*

$$F_t(L)\,\varepsilon_t = \frac{\phi_a}{\phi_c} a_t. \tag{65}$$

At the optimal monetary policy, the price level does not respond to aggregate technology shocks.

Proof. See Appendix C. ■

The reason for this result about optimal policy is the same as in the previous subsection: The response of the economy to aggregate technology shocks under perfect information is efficient and by offsetting the effect of aggregate technology shocks on the profit-maximizing price the central bank can replicate the perfect-information response of real variables to aggregate technology shocks. There is one difference to the previous subsection. At the optimal monetary policy, price setters in firms now devote no attention to aggregate technology because the profit-maximizing price does not respond to aggregate technology shocks.

7 Optimal monetary policy response to markup shocks

In this section, we derive the optimal monetary policy response to markup shocks. Our main result is the following. Complete price stabilization in response to markup shocks is never optimal in the model with an exogenous information structure, whereas complete price stabilization in response to markup shocks is always optimal in the model with an endogenous information structure.

For ease of exposition, we assume in this section that there are no aggregate technology shocks. This assumption simplifies the notation in Propositions 4, 5 and 6, and has no impact on the optimal monetary policy response to markup shocks.

7.1 Exogenous information structure

In the model with an exogenous information structure, the optimal monetary policy response to markup shocks in the case of an i.i.d. desired markup is given by the following proposition.

Proposition 4 *(Exogenous signal precision) Consider the Ramsey problem (42)-(56), where the variances of noise* σ_η^2 *and* σ_ζ^2 *are exogenous. Suppose* $\sigma_\nu^2 > 0$*,* $\rho_\lambda = 0$ *and* $\sigma_\varepsilon^2 = a_{-1} = 0$*. Consider policies of the form* $G_t(L)\,\nu_t = g_0\nu_t$ *and equilibria of the form* $p_t = \theta\lambda_t$*. The unique equilibrium at any monetary policy* $g_0 \in \mathbb{R}$ *is*

$$p_t = \frac{\phi_c g_0 + \phi_\lambda}{\phi_c + \frac{\sigma_\zeta^2}{\sigma_\lambda^2}}\lambda_t, \tag{66}$$

$$c_t = \frac{\frac{\sigma_\zeta^2}{\sigma_\lambda^2}g_0 - \phi_\lambda}{\phi_c + \frac{\sigma_\zeta^2}{\sigma_\lambda^2}}\lambda_t, \tag{67}$$

$$p_{i,t} - p_t = \frac{\phi_c g_0 + \phi_\lambda}{\phi_c + \frac{\sigma_\zeta^2}{\sigma_\lambda^2}}\zeta_{i,t}. \tag{68}$$

Furthermore, if $\sigma_\zeta^2 > 0$*, the unique optimal monetary policy* $g_0 \in \mathbb{R}$ *is*

$$g_0^* = \frac{(1 - \delta\phi_c)\,\phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2} + \delta\phi_c^2}. \tag{69}$$

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

Proof. See Appendix D. ■

The main result in Proposition 4 is that in the model with an exogenous information structure and an i.i.d. desired markup, 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 change the monetary policy instrument in response to markup shocks (i.e., $g_0 = 0$). In this case, a positive markup shock (i.e., a shock that raises the desired markup) increases the profit-maximizing price. Price setters in firms therefore put a positive weight on their signals concerning the desired markup which causes inefficient price dispersion due to noise in the signal ("cross-sectional inefficiency"). Furthermore, the price level increases which - given the constant money supply - causes a fall in consumption ("aggregate inefficiency"). To reduce inefficient price dispersion, the central bank can counteract the effect of a positive markup shock on the profit-maximizing price with a contractionary monetary policy (i.e., by lowering the money supply). The profit-maximizing price then increases by less in response to a positive markup shock, implying that price setters in firms put less weight on their noisy signals concerning the desired markup, which reduces inefficient price dispersion. Unfortunately, the contractionary monetary policy amplifies the fall in consumption after a positive markup shock. Thus, reducing cross-sectional inefficiency increases aggregate inefficiency. There exists a trade-off between cross-sectional inefficiency and aggregate inefficiency. Moreover, 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. Formally, substituting the optimal monetary policy (69) into equations (66) and (68) yields

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

The existence of a trade-off between cross-sectional inefficiency and aggregate inefficiency in the presence of markup shocks and the result that complete price stabilization in response to markup shocks is not optimal are classic results in monetary economics. These results hold in a variety of other models. See, for example, Woodford (2003), Gali (2008), Adam (2007), and Ball, Mankiw and Reis (2005).

We now turn to the case of an autocorrelated desired markup. We want to know whether complete price stabilization in response to markup shocks is still suboptimal when the desired markup is autocorrelated. When $\rho_\lambda > 0$ we solve the Ramsey problem (42)-(56) 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.[9] Following the procedure in Woodford (2002), one can then compute an exact linear rational expectations equilibrium of the model (43)-(56) *for a given monetary policy* by solving a Riccati equation. We then run a numerical optimization routine to obtain the optimal monetary policy.

[9] 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.

Figure 1 shows the optimal monetary policy response to a markup shock in the model with an exogenous signal precision for the following parameter values: $\beta = 0.99$, $\gamma = \psi = 1$, $\alpha = (2/3)$, $\Lambda = (1/4)$, $\sigma_\nu = 0.2$, and $\sigma_\zeta = 0.4$. The upper panel of Figure 1 shows optimal monetary policy when $\rho_\lambda = 0$. The lower panel of Figure 1 shows optimal monetary policy when $\rho_\lambda = 0.9$. The value for β is a standard value for a quarterly model. The values for γ, ψ and α are standard values in the business cycle literature. The value for Λ implies a steady-state price elasticity of demand of five. This value for the price elasticity of demand is half way between the value of three used by Midrigan (2010) and the value of seven used by Golosov and Lucas (2007).[10] The values $\rho_\lambda = 0.9$ and $\sigma_\nu = 0.2$ are within the range of estimates of markup shocks in the New Keynesian literature. For comparison, Figure 1 also shows the optimal monetary policy in the Calvo model. In the Calvo model, decision-makers in firms have perfect information and in every period each firm can adjust its price with an exogenous probability. Following Nakamura and Steinsson (2008), we set the fraction of firms that can adjust their price in a quarter equal to 0.4. Finally, the standard deviation of noise, σ_ζ, is set such that the model with an exogenous signal precision and the Calvo model yield the same response of the price level to a markup shock when the component of the profit-maximizing price driven by markup shocks is a random walk. The idea is that we want to compare the model with an exogenous signal precision and the Calvo model for parameter values that imply the same degree of stickiness of the price level. All impulse responses are to a positive one standard deviation markup shock. A response equal to one means a one percent deviation from the non-stochastic steady state. Time is measured in quarters along the horizontal axis.

The main result is the following. In the model with an exogenous signal precision, complete price stabilization in response to markup shocks is still suboptimal when $\rho_\lambda > 0$. At the optimal monetary policy, the price level strictly increases on impact of a positive markup shock. Figure 1 shows this result for our benchmark parameter values. We solved the Ramsey problem (42)-(56) for many sets of parameter values with $\rho_\lambda > 0$ and we always obtained this result.

Furthermore, comparing the optimal monetary policy response to markup shocks in the model with an exogenous signal precision and in the Calvo model, we obtain the following results. First, the optimal monetary policy is qualitatively similar in these two models: At the optimal monetary policy, the price level strictly increases on impact of a positive markup shock, the consumption

[10] A price elasticity of demand of five is towards the upper end of estimates in the empirical IO literature.

level strictly falls on impact of a positive markup shock, and there is inefficient price dispersion. Second, whether optimal monetary policy is also quantitatively similar in these two models depends on parameter values. When $\rho_\lambda = 0.9$ the impulse responses at the optimal monetary policy are very similar in the model with an exogenous signal precision and in the Calvo model. By contrast, when $\rho_\lambda = 0$ the optimal monetary policy in the model with an exogenous signal precision is to respond to the markup shock only in the period of the shock, whereas the optimal monetary policy in the Calvo model is to respond to the markup shock also in the periods after the shock.

To summarize, in the model with an exogenous signal precision, there exists a trade-off between reducing inefficient price dispersion and reducing inefficient consumption variance and complete price stabilization in response to markup shocks is not optimal. Furthermore, qualitatively optimal policy is similar to optimal policy in the Calvo model.

7.2 Endogenous information structure

In the model with an endogenous information structure, the optimal monetary policy response to markup shocks in the case of an i.i.d. desired markup is given by the following two propositions. Proposition 5 treats the case of $\phi_c \in \left[\frac{1}{2}, \infty\right)$ and Proposition 6 treats the case of $\phi_c \in \left(0, \frac{1}{2}\right)$.

Proposition 5 *(Endogenous signal precision) Consider the Ramsey problem (42)-(60), where the signal precisions* $\left(1/\sigma_\eta^2\right)$ *and* $\left(1/\sigma_\zeta^2\right)$ *are given by the solution to problem (57)-(60). Suppose that* $\mu > 0$, $\sigma_\nu^2 > 0$, $\rho_\lambda = 0$ *and* $\sigma_\varepsilon^2 = a_{-1} = 0$. *Consider policies of the form* $G_t(L)\,\nu_t = g_0 \nu_t$ *and equilibria of the form* $p_t = \theta\lambda_t$. *Define*

$$b \equiv \sqrt{\frac{\omega\left(\phi_c g_0 + \phi_\lambda\right)^2 \sigma_\lambda^2 \ln(2)}{\mu}}. \tag{70}$$

First, we characterize the set of equilibria at a given monetary policy $g_0 \in \mathbb{R}$. *Let* κ^* *denote the equilibrium attention devoted to the desired markup. If and only if* $b \leq 1$, *there exists an equilibrium with*

$$\kappa^* = 0. \tag{71}$$

If and only if $\phi_c \in \left(0, \frac{1}{2}\right]$ *and* $b \in \left[\sqrt{4\phi_c\left(1-\phi_c\right)}, 1\right]$, *there exists an equilibrium with*

$$\kappa^* = \log_2\left(\frac{b - \sqrt{b^2 - 4\phi_c\left(1-\phi_c\right)}}{2\phi_c}\right). \tag{72}$$

If and only if either $\phi_c \in \left(0, \frac{1}{2}\right]$ and $b \geq \sqrt{4\phi_c (1-\phi_c)}$ or $\phi_c > \frac{1}{2}$ and $b \geq 1$, there exists an equilibrium with

$$\kappa^* = \log_2 \left(\frac{b + \sqrt{b^2 - 4\phi_c (1-\phi_c)}}{2\phi_c} \right). \tag{73}$$

The equilibrium price level, consumption level and price dispersion are given by

$$p_t = \frac{(\phi_c g_0 + \phi_\lambda)\left(1 - 2^{-2\kappa^*}\right)}{1 - (1-\phi_c)\left(1 - 2^{-2\kappa^*}\right)} \lambda_t, \tag{74}$$

$$c_t = \left[g_0 - \frac{(\phi_c g_0 + \phi_\lambda)\left(1 - 2^{-2\kappa^*}\right)}{1 - (1-\phi_c)\left(1 - 2^{-2\kappa^*}\right)} \right] \lambda_t, \tag{75}$$

and

$$E\left[(p_{i,t} - p_t)^2\right] = \frac{\frac{\mu}{\omega}}{\ln(2)} \left(1 - 2^{-2\kappa^*}\right). \tag{76}$$

Second, we characterize optimal monetary policy. If $\phi_c \in \left[\frac{1}{2}, \infty\right)$, there exists a unique equilibrium for any monetary policy $g_0 \in \mathbb{R}$ and the unique optimal monetary policy is

$$g_0^* = \begin{cases} 0 & \text{if } \frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} \leq 1 \\ -\frac{\phi_\lambda}{\phi_c} + \frac{1}{\phi_c}\sqrt{\frac{\mu}{\omega \sigma_\lambda^2 \ln(2)}} & \text{if } \frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} > 1 \end{cases}. \tag{77}$$

At the optimal monetary policy, price setters in firms pay no attention to the desired markup (i.e., $\kappa^* = 0$), the price level does not respond to markup shocks, and there is no inefficient price dispersion.

Proof. See Appendix E. ■

Proposition 5 states that in the model with an endogenous information structure, complete price stabilization in response to markup shocks is optimal when $\mu > 0$, $\rho_\lambda = 0$ and $\phi_c \in \left[\frac{1}{2}, \infty\right)$. The condition $\mu > 0$ means that there is some cost of paying attention to the desired markup. This cost can be arbitrarily small. The condition $\rho_\lambda = 0$ means that the desired markup follows a white noise process. The condition $\phi_c \in \left[\frac{1}{2}, \infty\right)$ means that strategic complementarity in price setting is not strong enough for multiple equilibria to arise. Below we show analytically that in the model with an endogenous information structure complete price stabilization in response to markup shocks is also optimal when $\phi_c \in \left(0, \frac{1}{2}\right)$, $\mu > 0$ and $\rho_\lambda = 0$. Hence, when $\mu > 0$ and $\rho_\lambda = 0$ complete price stabilization in response to markup shocks is *always* optimal in the model with an *endogenous information structure*. This result is in the starkest possible contrast to Proposition 4 stating that when $\sigma_\zeta^2 > 0$ and $\rho_\lambda = 0$ complete price stabilization in response to markup shocks is *never* optimal in the model with an *exogenous information structure*.

To understand Proposition 5, let us first focus on the optimal allocation of attention by price setters in firms. This is the new feature in the model with an endogenous information structure. The profit-maximizing price of good i in period t equals

$$\begin{aligned} p_{i,t}^* &= (1-\phi_c)\,p_t + \phi_c m_t - \phi_a a_t + \phi_\lambda \lambda_t \\ &= \left[(1-\phi_c)\,\theta + \phi_c g_0 + \phi_\lambda\right]\lambda_t . \end{aligned}$$

A price setter's optimal amount of attention devoted to the desired markup λ_t equals

$$\kappa^* = \begin{cases} \frac{1}{2}\log_2\left(\frac{\omega[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]^2\sigma_\lambda^2\ln(2)}{\mu}\right) & \text{if } \frac{\omega[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]^2\sigma_\lambda^2\ln(2)}{\mu} \geq 1 \\ 0 & \text{otherwise} \end{cases}. \tag{78}$$

The ratio $\frac{\omega[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]^2\sigma_\lambda^2\ln(2)}{\mu}$ is the marginal benefit of devoting attention to the desired markup when no attention is devoted to the desired markup (i.e., $\kappa = 0$) divided by the marginal cost of devoting attention to the desired markup. If this ratio exceeds one, the price setter in a firm devotes some attention to the desired markup. If this ratio increases, the price setter devotes more attention to the desired markup. The benefit of paying attention to the desired markup depends on $\left[(1-\phi_c)\,\theta + \phi_c g_0 + \phi_\lambda\right]^2 \sigma_\lambda^2$ which is the variance of the profit-maximizing price due to the desired markup. This variance depends on the behavior of other firms through θ and on monetary policy through g_0. As pointed out by Maćkowiak and Wiederholt (2009) and Hellwig and Veldkamp (2009), strategic complementarity in price setting leads to strategic complementarity in the allocation of attention. When other firms are paying more attention to the desired markup, the price level responds more to the desired markup, which in the case of $(1-\phi_c) > 0$ raises the incentive for an individual firm to pay attention to the desired markup. For this reason, multiple equilibria can in principle arise. However, when $\phi_c \in \left[\frac{1}{2}, \infty\right)$, strategic complementarity in price setting is not strong enough for multiple equilibria to arise. In this case, if the compound parameter b defined by equation (70) is below one, the unique equilibrium attention is given by equation (71); and if the compound parameter b is above one, the unique equilibrium attention is given by equation (73). To illustrate this result, the upper panel of Figure 2 shows equilibrium attention as a function of the compound parameter b for $\phi_c = (1/2)$. By contrast, when $\phi_c \in \left(0, \frac{1}{2}\right)$, strategic complementarity in price setting is strong enough for multiple equilibria to arise at some values of the compound parameter b. To illustrate this result, the lower panel of Figure 2 shows equilibrium attention as a function of the compound parameter b for $\phi_c = (1/4)$.

Let us now turn to optimal monetary policy. Proposition 5 specifies optimal monetary policy when $\phi_c \in \left[\frac{1}{2}, \infty\right)$, that is, when there exists a unique equilibrium for any monetary policy $g_0 \in \mathbb{R}$. First, consider the case of $\frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} \leq 1$. In this case, if the central bank does not respond to markup shocks (i.e., $g_0 = 0$), decision-makers in firms pay no attention to the desired markup because the marginal benefit of devoting attention to random variation in the desired markup is smaller than the marginal cost of devoting attention to random variation in the desired markup. Furthermore, when neither the central bank nor firms respond to markup shocks, markup shocks create no inefficiencies. Hence, in the case of $\frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} \leq 1$, a monetary policy of no response to markup shocks implements the efficient allocation and is thus the optimal monetary policy.

Second, consider the case of $\frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} > 1$. In this case, if the central bank does not respond to markup shocks (i.e., $g_0 = 0$), decision-makers in firms pay attention to the desired markup because the marginal benefit of devoting attention to random variation in the desired markup exceeds the marginal cost of devoting attention to random variation in the desired markup. Hence, in the case of $\frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} > 1$, if the central bank does not respond to markup shocks, the price level increases after a positive markup shock, the consumption level falls after a positive markup shock, and there is inefficient price dispersion caused by the noise in the signal concerning the desired markup. Next, suppose instead that the central bank counteracts the effect of a positive markup shock on the profit-maximizing price with a contractionary monetary policy (i.e., $g_0 < 0$). There are two effects that are already present in the model with an exogenous information structure: (i) the profit-maximizing price increases by less after a positive markup shock, implying that decision-makers in firms put less weight on their noisy signals concerning the desired markup, which reduces inefficient price dispersion; and (ii) the contractionary monetary policy by itself amplifies the fall in consumption after a positive markup shock. In addition, there is a new effect that is only present in the model with an endogenous information structure: When the profit-maximizing price increases by less after a positive markup shock, the variance of the profit-maximizing price due to markup shocks falls and thus decision-makers in firms decide to pay less attention to the desired markup. The price level therefore increases by less after a positive markup shock, which by itself mutes the fall in consumption after a positive markup shock. It turns out that this new effect on consumption which is only present in the model with an endogenous information structure dominates the old effect on consumption which is already present in the model with an exogenous information structure for all parameter values. Thus, so long as decision-makers in firms pay some attention to the desired markup, the central bank can reduce both inefficient price dispersion and inefficient consumption variance by counteracting the effect of a markup shock on the profit-maximizing price more strongly. Hence, the classic trade-off between reducing cross-sectional inefficiency and reducing aggregate inefficiency disappears. It is now straightforward to derive optimal monetary policy. So long as price setters in firms pay some attention to the desired markup, the central bank can reduce both inefficient price dispersion and inefficient consumption variance by counteracting markup shocks more strongly. Once price setters in firms pay no attention to the desired markup, inefficient price dispersion equals zero and reducing the money supply even more strongly after a positive markup shock only increases the fall in consumption after a positive markup shock. Hence, the unique optimal monetary policy is the one that makes decision-makers in firms just pay no attention to the desired markup. The lower part of equation (77) specifies this policy. At the optimal monetary policy, price setters in firms pay no attention to random variation in the desired markup (i.e., $\kappa^* = 0$) and therefore the price level does not respond to markup shocks. Complete price stabilization in response to markup shocks is optimal.

Figure 3 illustrates Proposition 5 for the following parameter values: $\gamma = \psi = 1$, $\alpha = (2/3)$, $\Lambda = (1/4)$, $\rho_\lambda = 0$ and $\sigma_\nu^2 = \left[(0.2)^2 / \left(1 - (0.9)^2\right)\right]$. Figure 3 depicts the optimal monetary policy, the amount of attention that price setters devote to the desired markup at the optimal policy, and the loss in welfare due to markup shocks at the optimal policy for different values of (μ/ω). Recall that $\mu > 0$ is the marginal cost of devoting attention for the decision-maker in a firm and $\omega > 0$ is the constant in the price setters' objective (57). The optimal monetary policy makes decision-makers in firms pay no attention to the desired markup independent of the value of (μ/ω).

Proposition 5 specifies optimal monetary policy when $\phi_c \in \left[\frac{1}{2}, \infty\right)$, that is, when strategic complementarity in price setting is not large enough for multiple equilibria to arise. Proposition 6 specifies optimal monetary policy when $\phi_c \in \left(0, \frac{1}{2}\right)$, that is, when strategic complementarity in price setting is large enough for multiple equilibria to arise at some monetary 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. The most common assumption in the literature is that central banks are very adverse to policies that yield multiple equilibria. Therefore, we assume that the central bank aims to implement the best policy among all those monetary policies $g_0 \in \mathbb{R}$ that yield a unique equilibrium.

Proposition 6 *(Endogenous signal precision) Consider the Ramsey problem (42)-(60), where the signal precisions* $\left(1/\sigma_\eta^2\right)$ *and* $\left(1/\sigma_\zeta^2\right)$ *are given by the solution to problem (57)-(60). Suppose that* $\mu > 0$, $\sigma_\nu^2 > 0$, $\rho_\lambda = 0$ *and* $\sigma_\varepsilon^2 = a_{-1} = 0$. *Consider policies of the form* $G_t(L)\nu_t = g_0\nu_t$ *and equilibria of the form* $p_t = \theta\lambda_t$. *If* $\phi_c \in \left(0, \frac{1}{2}\right)$, *there exist multiple equilibria for all* $g_0 \in [\hat{g}_0, \bar{g}_0]$ *where*

$$\hat{g}_0 = -\frac{\phi_\lambda}{\phi_c} + \frac{\sqrt{4\phi_c(1-\phi_c)}}{\phi_c}\sqrt{\frac{\mu}{\omega\sigma_\lambda^2 \ln(2)}}, \tag{79}$$

and

$$\bar{g}_0 = -\frac{\phi_\lambda}{\phi_c} + \frac{1}{\phi_c}\sqrt{\frac{\mu}{\omega\sigma_\lambda^2 \ln(2)}}. \tag{80}$$

If $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu} < 4\phi_c(1-\phi_c)$, *the best policy among all* $g_0 \in \mathbb{R}$ *that yield a unique equilibrium is* $g_0^* = 0$. *If* $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$, *the best policy among all* $g_0 \in \mathbb{R}$ *that yield a unique equilibrium is a* g_0 *marginally below* $\hat{g}_0$. *At this policy, price setters in firms pay no attention to the desired markup, the price level does not respond to markup shocks, and there is no inefficient price dispersion.*

Proof. See Appendix E. ■

The main result in Proposition 6 is that in the model with an endogenous information structure complete price stabilization in response to markup shocks is also optimal when $\phi_c \in \left(0, \frac{1}{2}\right)$. To understand this result, note the following. When $g_0 \in [\hat{g}_0, \bar{g}_0]$, the compound parameter b governing the benefit to the cost of paying attention to markup shocks lies in the interval $\left[\sqrt{4\phi_c(1-\phi_c)}, 1\right]$. Furthermore, the first half of Proposition 5 states that, if $\phi_c \in \left(0, \frac{1}{2}\right)$ and $b \in \left[\sqrt{4\phi_c(1-\phi_c)}, 1\right]$, then multiple equilibria arise. See Figure 2 for an illustration. Hence, if $\phi_c \in \left(0, \frac{1}{2}\right)$, the central bank has to choose $g_0 \notin [\hat{g}_0, \bar{g}_0]$ to avoid multiple equilibria. Next, think about optimal monetary policy. When $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu} < 4\phi_c(1-\phi_c)$ we have $\hat{g}_0 > 0$. Thus, at the policy $g_0 = 0$, $\kappa^* = 0$ is the unique equilibrium. When the central bank and firms do not respond to markup shocks, those shocks create no inefficiencies. Hence, in the case of $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu} < 4\phi_c(1-\phi_c)$, a monetary policy of no response to markup shocks is the optimal monetary policy. By contrast, when $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$ we have $\hat{g}_0 \leq 0$. Thus, at the policy $g_0 = 0$, $\kappa^* = 0$ is *not* the unique equilibrium. To understand optimal monetary policy in this case, consider the lower panel of Figure 2. When $g_0 > \bar{g}_0$ and thus $b > 1$, both price dispersion and consumption variance fall when the central bank *reduces* g_0 for the same reasons given below Proposition 5. When $g_0 < \hat{g}_0$ and thus $b < \sqrt{4\phi_c (1 - \phi_c)}$, consumption variance falls when the central bank *increases* g_0. Finally, the no attention equilibrium at $g_0 = \hat{g}_0$ strictly dominates the high positive attention equilibrium at $g_0 = \bar{g}_0$. Hence, the best policy among all policies that yield a unique equilibrium is a g_0 marginally below $\hat{g}_0$. At this policy, price setters in firms pay no attention to the desired markup and therefore the price level does not respond to markup shocks. Complete price stabilization in response to markup shocks is optimal.

Finally, we consider the case of an autocorrelated desired markup. When $\rho_\lambda > 0$ we solve the Ramsey problem (42)-(60) numerically. In particular, we restrict the infinite-order lag polynomial $G_t (L)$ in equation (52) to be the lag polynomial of an ARMA(2,2) process.[11] Following the procedure in Woodford (2002), one can then compute an exact linear rational expectations equilibrium of the model (43)-(56) for a given monetary policy and for given signal precision $(1/\sigma_\zeta^2)$. Furthermore, for a given law of motion for the endogenous variables, one can solve the attention problem (57)-(60). 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. Finally, we solve for the optimal monetary policy both by using an optimization routine and by evaluating the central bank's objective for different policies on a fine grid. Figure 4 depicts the optimal monetary policy for the following parameter values: $\beta = 0.99$, $\gamma = 1$, $\psi = 0$, $\alpha = 1$, $\Lambda = (1/4)$, $\rho_\lambda = 0.9$, $\sigma_\nu = 0.2$, and $(\mu/\omega) = 10^{-4}$. The figure shows the impulse responses of the money supply, the consumption level, the price level, and the profit-maximizing price to a one standard deviation positive markup shock. The optimal monetary policy is to reduce the money supply on impact of a positive markup shock so as to counteract the effect of the markup shock on the profit-maximizing price. At the optimal monetary policy, price setters in firms devote no attention to random variation in the desired markup (i.e., $\kappa^* = 0$) and therefore the price level does not respond to markup shocks. Complete price stabilization in response to markup shocks is optimal. We solved the Ramsey problem (42)-(60) for many sets of parameter values with $\rho_\lambda > 0$ and we always obtained this result.[12]

[11] See Footnote 9.

[12] We solved the Ramsey problem (42)-(60) for values of (μ/ω) between 10^{-5} and 1. We also solved the Ramsey problem (42)-(60) for $\psi = 1$ and $\alpha = (2/3)$.

8 Additional results and robustness of the main results

In this section, we present additional results for the model with an endogenous information structure (Sections 8.1-8.2). Furthermore, we show that the main results for the model with an endogenous information structure are robust to several modifications of the model (Sections 8.3-8.6). The most important results are the following. First, the result that complete price stabilization is optimal in response to markup shocks extends to other shocks that cause inefficient fluctuations under perfect information. Second, optimal monetary policy remains the same when decision-makers in firms can decide to receive signals concerning any linear combination of a_t and λ_t (e.g., signals concerning endogenous variables).

8.1 Welfare at the optimal monetary policy

How does welfare at the optimal monetary policy depend on price setters' marginal cost of paying attention to aggregate conditions?

In the model with an endogenous information structure and $\rho_\lambda = 0$, the value of the central bank's objective (42) at the optimal monetary policy specified in Propositions 3, 5 and 6 equals

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

where g_0^* is given by Proposition 5 if $\phi_c \in \left[\frac{1}{2}, \infty\right)$ and g_0^* is given by Proposition 6 if $\phi_c \in \left(0, \frac{1}{2}\right)$. Recall that the absolute value of g_0^* is weakly decreasing in μ. See Figure 3 for an illustration. Therefore, welfare is weakly increasing in price setters' marginal cost of paying attention to aggregate conditions. The intuition is simple. When price setters' marginal cost of paying attention to markup shocks is smaller, the central bank has to counteract markup shocks more strongly to discourage price setters from paying attention to these shocks that cause inefficient fluctuations. Hence, in the model with an endogenous information structure, easier access to information concerning markup shocks reduces welfare. This result extends to other shocks that cause inefficient fluctuations under perfect information. See Section 8.4.

8.2 Gain from commitment

In the model with an endogenous information structure, there is a gain from commitment to a policy rule when there are markup shocks and the central bank has to counteract the effect of markup shocks on the profit-maximizing price to discourage price setters from paying attention to these shocks ($g_0^* < 0$). In this case, only when the central bank can commit, price setters can trust the central bank that paying no attention to aggregate conditions is optimal. In the Calvo model, there is also a gain from commitment to a monetary policy rule when there are markup shocks. In particular, when the central bank can commit, price setters can trust the central bank that they will not be exploited during the period where they cannot adjust their prices. We find it interesting that the gain from commitment in the rational inattention model is of different nature than the gain from commitment in the Calvo model.

8.3 More general signal structure

We have so far assumed that paying attention to aggregate technology and paying attention to the desired markup are independent activities. Formally, in equation (17) we assume that the price setter in firm i receives independent signals concerning aggregate technology and the desired markup. We now relax this assumption. We now assume that the price setter in firm i can pay attention to any variable that is a linear combination of a_t and λ_t. Formally, the signal that the price setter in firm i receives in period t can be any signal of the form

$$s_{i,t} = \xi_a a_t + \xi_\lambda \lambda_t + \zeta_{i,t}. \tag{81}$$

The decision-maker chooses both the coefficients $(\xi_a, \xi_\lambda) \in \mathbb{R}^2$ and the variance of noise $\sigma_\zeta^2 \in \mathbb{R}_+$.[13] The choice of (ξ_a, ξ_λ) can be interpreted as the choice of which variable to pay attention to. More precisely, when $\rho_a = \rho_\lambda = 0$ all endogenous variables are just linear functions of a_t and λ_t and thus equation (81) implies that the price setter in a firm can pay attention to any endogenous variable. For example, the price setter can choose the coefficients ξ_a and ξ_λ such that the linear combination $\xi_a a_t + \xi_\lambda \lambda_t$ equals the equilibrium price level or equilibrium output (i.e., the price setter can choose $s_{i,t} = p_t + \zeta_{i,t}$ or $s_{i,t} = c_t + \zeta_{i,t}$). The choice of σ_ζ^2 can then be interpreted as the choice of how much attention to devote to the price level or output.

For the signal structure (17), equation (20) reduces to equation (59). For the signal structure (81), this is not the case and we have to work with the original equation (20). Thus, in the new model with an endogenous information structure, the Ramsey problem (42)-(60) changes as follows:

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

Equation (81) replaces equation (49), equation (20) replaces equation (59), and the price setter in a firm chooses $\left(\xi_a, \xi_\lambda, 1/\sigma_\zeta^2\right)$ instead of $\left(1/\sigma_\eta^2, 1/\sigma_\zeta^2\right)$.

In the case of $\rho_a = \rho_\lambda = 0$, we can solve this new Ramsey problem analytically. In particular, the optimal monetary policy is again the monetary policy specified in Propositions 3, 5 and 6. The reason is quite simple. Consider first the attention decision of the decision-maker in a firm who has to set a price. When $\rho_a = \rho_\lambda = 0$, the decision-maker chooses to pay attention directly to the profit-maximizing price, that is, the optimal choice of (ξ_a, ξ_λ) is the (ξ_a, ξ_λ) with the property that $\xi_a a_t + \xi_\lambda \lambda_t$ equals the equilibrium profit-maximizing price.[14] Furthermore, the optimal amount of attention devoted to the profit-maximizing price equals

$$\kappa^* = \begin{cases} \frac{1}{2}\log_2\left(\frac{\omega\sigma_{p^*}^2 \ln(2)}{\mu}\right) & \text{if } \frac{\omega\sigma_{p^*}^2 \ln(2)}{\mu} \geq 1 \\ 0 & \text{otherwise} \end{cases}, \tag{82}$$

where $\sigma_{p^*}^2$ denotes the variance of the profit-maximizing price. The signal-to-noise ratio of the signal (81) then equals

$$\frac{\sigma_{p^*}^2}{\sigma_\zeta^2} = 2^{2\kappa^*} - 1.$$

The price set by the decision-maker in firm 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 3. 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 κ^*. Moreover, a small κ^* is good because then prices respond less to markup shocks. For these reasons, the optimal monetary policy response to aggregate technology shocks is the monetary policy specified in Proposition 3. Second, once the profit-maximizing price does not respond to aggregate technology shocks, equation (82) reduces to equation (78). In other words, the firms' optimal allocation of attention is exactly the same as in the model with no aggregate technology shocks. For this reason, the optimal monetary policy response to markup shocks is the monetary policy specified in Propositions 5 and 6.

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

8.4 More general shocks

We now show that the results for aggregate technology shocks and markup shocks presented in Sections 6-7 extend to a much larger class of shocks. The results for aggregate technology shocks extend to other shocks that cause efficient fluctuations under perfect information. The results for markup shocks extend to other shocks that cause inefficient fluctuations under perfect information.

We begin by introducing a general exogenous aggregate variable. This variable, denoted z_t, may affect efficient composite consumption and/or the profit-maximizing price:

$$c_t^* = \varphi z_t, \tag{83}$$

and

$$p_{i,t}^* = p_t + \phi_c c_t + \phi_z z_t, \tag{84}$$

where $\varphi \in \mathbb{R}$ and $\phi_z \in \mathbb{R}_{++}$. The variable z_t is assumed to follow a stationary Gaussian first-order autoregressive process. One example of the variable z_t is aggregate technology in the model given in Section 2. In this case, $z_t = -a_t$, $\varphi = -\frac{\phi_a}{\phi_c}$ and $\phi_z = \phi_a$. Another example of the variable z_t is the desired markup in the model given in Section 2. In that case, $z_t = \lambda_t$, $\varphi = 0$ and $\phi_z = \phi_\lambda$. Apart from introducing a general aggregate exogenous variable the Ramsey problem (42)-(60) remains unchanged.[15] In the new Ramsey problem, equations (83) and (84) replace equations (43) and (47), and $s_{i,t} = z_t + \zeta_{i,t}$ and $\kappa = \frac{1}{2}\log_2\left(\frac{\sigma^2_{z|t-1}}{\sigma^2_{z|t}}\right)$ replace equations (49) and (59). Price setters choose the precision of their signals concerning z_t knowing that a more precise signal requires more attention. The question is: What is the optimal monetary policy response to an innovation in z_t?

If price setters in firms had perfect information, each firm would set the profit-maximizing price, implying that

$$c_t = -\frac{\phi_z}{\phi_c} z_t,$$

and

$$p_{i,t} - p_t = 0.$$

Hence, the response of the economy to an innovation in the variable z_t under perfect information is efficient if and only if $\varphi = -\frac{\phi_z}{\phi_c}$.

[15] For ease of exposition, we now assume in equations (83)-(84) that there is only one exogenous variable.

Turning to optimal policy, consider the model with an exogenous information structure. It is straightforward to extend Proposition 4 from the variable λ_t to the more general variable z_t. The new version of equation (69) reads

$$g_0^* = \frac{(1-\delta\phi_c)\,\phi_z + \varphi\left(\phi_c + \frac{\sigma_\zeta^2}{\sigma_z^2}\right)}{\frac{\sigma_\zeta^2}{\sigma_z^2} + \delta\phi_c^2}.$$

The price level at the optimal monetary policy equals

$$p_t = \frac{\phi_z + \varphi\phi_c}{\frac{\sigma_\zeta^2}{\sigma_z^2} + \delta\phi_c^2} z_t.$$

Hence, in the model with an exogenous information structure and $\rho_z = 0$, complete price stabilization in response to an innovation in z_t is optimal if and only if the response of the economy to the shock under perfect information is efficient (i.e., $\varphi = -\frac{\phi_z}{\phi_c}$).

Next consider the model with an endogenous information structure. Here we distinguish three cases: (1) $\varphi = -\frac{\phi_z}{\phi_c}$, (2) $\varphi > -\frac{\phi_z}{\phi_c}$, and (3) $\varphi < -\frac{\phi_z}{\phi_c}$. Case (1) means that the response of the economy to an innovation in z_t under perfect information is efficient. In this case, the reasoning of Section 6 applies. By setting $m_t = -\frac{\phi_z}{\phi_c} z_t$ the central bank can replicate the response of the economy to the shock under perfect information and this response is efficient. Hence, the optimal monetary policy is $m_t = -\frac{\phi_z}{\phi_c} z_t$. Complete price stabilization in response to the shock is optimal. Case (2) means that the response of the economy to an innovation in z_t under perfect information is inefficient because the response is too large in magnitude ($0 \geq \varphi > -\frac{\phi_z}{\phi_c}$) or has the wrong sign ($\varphi > 0 > -\frac{\phi_z}{\phi_c}$). An example of a variable z_t with the property $\varphi > -\frac{\phi_z}{\phi_c}$ is the desired markup in the model given in Section 2 because under perfect information composite consumption responds to a markup shock while the efficient response of composite consumption to a markup shock equals zero. The proofs of Propositions 5 and 6 extend in a straightforward way from the desired markup to any variable z_t with the property $\varphi > -\frac{\phi_z}{\phi_c}$.[16] The beginning of the generalized version of Proposition 5 reads: $\varphi > -\frac{\phi_z}{\phi_c}$, $\sigma_z^2 > 0$, $\rho_z = 0$, $m_t = g_0 z_t$ and $p_t = \theta z_t$. The only changes in equations (70)-(76) are that ϕ_z, σ_z^2 and z_t replace ϕ_λ, σ_λ^2 and λ_t. The equation for optimal monetary policy becomes

$$g_0^* = \begin{cases} \varphi & \text{if } \frac{\omega(\phi_c\varphi+\phi_z)^2\sigma_z^2\ln(2)}{\mu} \leq 1 \\ -\frac{\phi_z}{\phi_c} + \frac{1}{\phi_c}\sqrt{\frac{\mu}{\omega\sigma_z^2\ln(2)}} & \text{if } \frac{\omega(\phi_c\varphi+\phi_z)^2\sigma_z^2\ln(2)}{\mu} > 1 \end{cases}.$$

[16]Statement and proof of the generalized version of Proposition 5 and of the generalized version of Proposition 6 are in the Technical Appendix to this paper which is available on our websites.

At the optimal monetary policy, price setters in firms pay no attention to the variable z_t and therefore the price level does not respond to an innovation in z_t. In other words, complete price stabilization is optimal in response to any variable z_t with the property $\varphi > -\frac{\phi_z}{\phi_c}$. Finally, case (3) means that the response of the economy to an innovation in z_t under perfect information is inefficient because the response is too small. The proofs of Propositions 5 and 6 do not extend in a straightforward way to variables that cause fluctuations under perfect information that are too small. Therefore, we do not know yet whether complete price stabilization in response to such shocks is optimal. This may or may not be the case.

To summarize, in the model with an exogenous information structure, complete price stabilization is optimal only in response to variables z_t that cause efficient fluctuations under perfect information (i.e., $\varphi = -\frac{\phi_z}{\phi_c}$). In the model with an endogenous information structure, complete price stabilization is optimal also in response to variables z_t that cause fluctuations under perfect information that are too large or have the wrong sign (i.e., $\varphi \geq -\frac{\phi_z}{\phi_c}$).[17] The optimality of complete price stabilization becomes a much more general result.

8.5 Interest rate rule

We now assume that the central bank commits to an interest rate rule instead of a money supply rule. Assuming that the central bank can commit to an interest rate rule of the form (11) instead of a money supply rule of the form (10) does not change optimal monetary policy. This is because the set of equilibria that the central bank can implement with an interest rate rule of the form (11) equals the set of equilibria that the central bank can implement with a money supply rule of the form (10). To see this, note the following. We so far did not use the log-linearized consumption Euler equation in the Ramsey problem (42)-(60) because in the case of a money supply rule this equation only determines the equilibrium nominal interest rate. Now take a law of motion of the economy that is an equilibrium law of motion under some money supply rule of the form (10). One can then compute the equilibrium law of motion for the nominal interest rate from the consumption Euler equation and the central bank can commit to this law of motion as an interest rate rule. Similarly,

[17] One example of a variable z_t with the property $\varphi > -\frac{\phi_z}{\phi_c}$ is a markup shock. Another example of a variable z_t with the property $\varphi > -\frac{\phi_z}{\phi_c}$ is an aggregate technology shock or a labor supply shock in an economy with a positive production or consumption externality.

take a law of motion of the economy that is an equilibrium law of motion under some interest rate rule of the form (11). One can then compute the equilibrium law of motion for the money supply from equation (44) and the central bank can commit to this law of motion as a money supply rule.[18]

8.6 Monetary transaction frictions

So far we have studied an economy without monetary transaction frictions. In equations (4)-(5) we assume that wage income can be transformed immediately into cash and cash can be used immediately to purchase goods. Therefore, the requirement that households need cash to purchase goods creates no distortions. Another interpretation of this economy is that this is a cashless economy à la Woodford (2003). To see this, note that substituting equation (5) into equation (4) yields the flow budget constraint of a cashless economy and the first-order conditions of the household are identical to the first-order conditions in a cashless economy. Furthermore, the monetary policy rule (10) can be interpreted as saying that the central bank commits to a law of motion for nominal spending. The central bank can implement this law of motion for nominal spending with a short-term nominal interest rate, as described in Section 8.5. Hence, the results presented up to here can be interpreted as the results for a cashless economy.

We now study an economy with transaction frictions. In particular, we assume that cash has to be held for one period before it can be used to purchase goods. Formally, we replace the flow budget constraint (4) and the cash-in-advance constraint (5) with

$$M_t + B_t = R_{t-1}B_{t-1} + W_t L_t + D_t - T_t + \left(M_{t-1} + X_t - \sum_{i=1}^{I} P_{i,t}C_{i,t}\right),$$

and

$$\sum_{i=1}^{I} P_{i,t}C_{i,t} = M_{t-1} + X_t.$$

Here $X_t = M_t^s - M_{t-1}^s$ is a nominal transfer and M_t denotes money demand which in equilibrium equals money supply (i.e., $M_t = M_t^s$). The new formulation of the flow budget constraint and the new formulation of the cash-in-advance constraint change one of the equations characterizing

[18] One issue that arises when the central bank commits to an interest rate rule of the form (11) instead of a money supply rule of the form (10) is the following. When the policy rule specifies the nominal interest rate as a function of exogenous events, the equilibrium is typically not unique. One way to address this issue is to allow the nominal interest rate to depend on endogenous variables and to allow the policy rule to differ on and off the equilibrium path.

equilibrium in the Ramsey problem (30)-(41). In particular, equation (33) for the optimal labor supply becomes

$$\frac{W_t}{P_t} = L_t^{\psi} C_t^{\gamma} \left(\beta E_t \left[\frac{C_{t+1}^{-\gamma} P_t}{C_t^{-\gamma} P_{t+1}} \right] \right)^{-1}.$$

This implies that equation (47) for the profit-maximizing price becomes

$$\begin{aligned} p_{i,t}^* &= p_t + \phi_c c_t - \phi_a a_t + \phi_\lambda \lambda_t \\ &\quad + \frac{1}{1 + \frac{1-\alpha}{\alpha} \frac{1+\Lambda}{\Lambda}} E_t \left[\gamma \left(c_{t+1} - c_t \right) + \left(p_{t+1} - p_t \right) \right]. \end{aligned}$$

In the new Ramsey problem for the economy with transaction frictions, we assume that the central bank corrects the distortion due to transaction frictions in the non-stochastic steady state by implementing the Friedman rule in the non-stochastic steady state. The money supply rule (41) becomes

$$\ln \left(M_t \right) = \ln \left(\beta \right) t + F_t \left(L \right) \varepsilon_t + G_t \left(L \right) \nu_t.$$

The variable m_t in equations (44) and (52) now denotes the log-deviation of the money supply from its deterministic trend β^t. When the desired markup is i.i.d. across time, one can again solve analytically for the equilibrium allocation as a function of monetary policy.[19] It is then straightforward to solve numerically for the optimal monetary policy. For our benchmark parameter values (i.e., $\gamma = \psi = 1$, $\alpha = (2/3)$, $\Lambda = (1/4)$ and $\sigma_\lambda^2 = \left[(0.2)^2 / \left(1 - (0.9)^2 \right) \right]$), we obtain the following results concerning optimal monetary policy in the economy with transaction frictions.[20] There is a threshold value for the ratio (μ/ω) above which complete price stabilization in response to markup shocks is optimal and below which complete price stabilization in response to markup shocks is no longer optimal. In particular, for values of (μ/ω) above the threshold, the optimal monetary policy is again to make price setters in firms pay no attention to the desired markup. When $\rho_\lambda = 0$ the threshold value for (μ/ω) equals $0.9 * 10^{-4}$ and when $\rho_\lambda = 0.9$ the threshold value for (μ/ω) is even lower. For comparison, note the following. Maćkowiak and Wiederholt (2010) solve a DSGE model with rational inattention and a Taylor rule and find that for a value of (μ/ω) between $1 * 10^{-4}$ and $2 * 10^{-4}$ the model matches various empirical impulse responses of prices to shocks. For this value of (μ/ω), complete price stabilization in response to markup shocks is optimal in our economy with markup shocks and transaction frictions.

[19] See Proposition 3 in the Technical Appendix to this paper which is available on our websites.

[20] Here σ_λ^2 denotes the variance of λ_t.

8.7 Information of the central bank

So far we did not model the information choice of the central bank. We simply assumed that the central bank has perfect information and can therefore implement the optimal monetary policy. We now study the central bank's benefit of learning aggregate conditions. If the central bank has no information about aggregate conditions, the central bank cannot respond to aggregate technology shocks and markup shocks. Figure 5 shows the central bank's benefit of learning aggregate conditions for the following parameter values: $\gamma = \psi = 1$, $\alpha = (2/3)$, $\Lambda = (1/4)$, $\rho_a = \rho_\lambda = 0$, $\sigma_a^2 = \left[(0.0085)^2 / \left(1 - (0.95)^2\right)\right]$ and $\sigma_\lambda^2 = \left[(0.2)^2 / \left(1 - (0.9)^2\right)\right]$.[21] In particular, the upper panel shows the loss in welfare (compared to the efficient allocation) in the case of the optimal monetary policy response to aggregate technology shocks and in the case of no policy response to aggregate technology shocks. The lower panel of Figure 5 shows the loss in welfare (compared to the efficient allocation) in the case of the optimal monetary policy response to markup shocks and in the case of no policy response to markup shocks. For values of (μ/ω) between $1 * 10^{-4}$ and $4 * 10^{-4}$, the per-period welfare gain from implementing the optimal monetary policy is quite large: about one third of a percent of steady state consumption.[22] Hence, for these values of (μ/ω), the central bank has a substantial incentive to become informed about aggregate conditions to implement the optimal monetary policy. At the same time, this optimal monetary policy makes the variance of the profit-maximizing price due to aggregate shocks sufficiently small such that price setters in firms pay no attention to aggregate conditions.[23]

Furthermore, the central bank does not literally have to know aggregate technology and the desired markup. Suppose that the central bank knows output, the output gap and the price level, where the output gap is defined as output minus efficient output. The central bank can then

[21] The parameter values $\rho_a = 0.95$ and $\sigma_\varepsilon = 0.0085$ are a standard calibration of the aggregate productivity process. We set $\rho_a = 0$ and $\sigma_a^2 = \left[(0.0085)^2 / \left(1 - (0.95)^2\right)\right]$ to make Figure 5 comparable to Figure 3. The figure looks similar for $\rho_a = 0.95$ and $\sigma_\varepsilon = 0.0085$, which again implies $\sigma_a^2 = \left[(0.0085)^2 / \left(1 - (0.95)^2\right)\right]$.

[22] Maćkowiak and Wiederholt (2010) solve a DSGE model with rational inattention and a Taylor rule. They find that for a value of (μ/ω) between $1 * 10^{-4}$ and $2 * 10^{-4}$ the model matches various empirical impulse responses of prices to shocks.

[23] In the welfare calculations, we are not taking into account that the optimal monetary policy has additional welfare benefits due to the fact that price setters do not have to pay attention to aggregate conditions and can thus focus on firm-specific conditions. Taking these additional welfare benefits into account would strengthen the case for complete price stabilization in response to aggregate shocks and would increase the welfare gain from optimal monetary policy.

implement a monetary policy that is arbitrarily close to the optimal monetary policy. The idea is simple: if the central bank encourages firms to pay a little bit of attention to the desired markup, the price level reveals the desired markup to the central bank and nevertheless the loss in welfare due to deviations from the optimal monetary policy can be made arbitrarily small. Formally, consider the case of $\rho_\lambda = 0$, $\phi_c \in \left[\frac{1}{2}, \infty\right)$, and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} > 1$.[24] Suppose that the central bank implements the following monetary policy

$$m_t = c_t^* + g_0 \frac{1 - (1 - \phi_c)\left(1 - 2^{-2\kappa^*}\right)}{(\phi_c g_0 + \phi_\lambda)\left(1 - 2^{-2\kappa^*}\right)} p_t. \tag{85}$$

Here c_t^* is efficient composite consumption, $g_0 \in (g_0^*, 0)$ where g_0^* is given by equation (77), κ^* is given by equations (70) and (73), and the ratio in front of the price level is the inverse of the ratio in equation (74). Since $g_0 > g_0^*$, price setters pay some attention to markup shocks and the price level reveals the desired markup to the central bank. At the same time, by making the difference $g_0 - g_0^*$ arbitrarily small, the central bank can approximate the optimal monetary policy arbitrarily well.

In the New Keynesian literature on optimal monetary policy, it is typically assumed that the central bank knows the output gap and inflation, where the output gap is defined as output minus efficient output. See for example Woodford (2003), Chapter 8, and Giannoni and Woodford (2010). Thus, when we assume that the central bank knows output, the output gap and the price level, we make essentially the same assumption about information of the central bank as the standard New Keynesian literature. Despite this, we obtain a markedly different result concerning optimal monetary policy: the optimal monetary policy is arbitrarily close to complete price stabilization in response to markup shocks and the only purpose of arbitrarily small fluctuations in the price level is to reveal the desired markup to the central bank.

We think it would be interesting to study optimal monetary policy when the central bank only observes noisy indicators of the output gap and inflation. Svensson and Woodford (2003, 2004) study this question in a New Keynesian model and Lorenzoni (2010) studies this question in a model with exogenous dispersed information. For studying this question, it will be useful to know what the central bank should do when the central bank has perfect information, which is the content of this paper.

[24] The case $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} > 1$ is the interesting case because in the case of $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \leq 1$ the central bank needs no information about the realization of the desired markup to implement the optimal policy. See Proposition 5.

9 Conclusion

This paper is the first paper studying a Ramsey optimal policy problem for an economy where decision-makers in firms choose how much attention they devote to aggregate conditions. Our main finding is that complete price stabilization is the optimal monetary policy *not only* in response to shocks that cause efficient fluctuations under perfect information *but also* in response to shocks that cause inefficient fluctuations under perfect information. At the optimal monetary policy, price setters in firms pay no attention to aggregate conditions.

Furthermore, we find that reducing price setters' marginal cost of paying attention to aggregate conditions reduces welfare, because then the central bank has to counteract the effect of markup shocks on the profit-maximizing price more strongly to discourage price setters in firms from paying attention to these shocks that cause inefficient fluctuations.

In addition, we find that there is a gain from commitment by the central bank, because only when the central bank can commit, price setters can trust the central bank that paying no attention to aggregate conditions is optimal.

A Proof of Proposition 1

First, we introduce notation. The function u is given by equation (24). Let x_t denote the vector of all arguments of the function u that are endogenous variables

$$x_t = \begin{pmatrix} c_t & \hat{c}_{1,t} & \cdots & \hat{c}_{I-1,t} \end{pmatrix}'.$$

Let z_t denote the vector of all arguments of the function u that are exogenous variables

$$z_t = \begin{pmatrix} a_t & \lambda_t \end{pmatrix}'.$$

Second, we compute a log-quadratic approximation to the period utility function (23) around the non-stochastic steady state. Let $\tilde{u}$ denote the second-order Taylor approximation to the function u at the origin. We have

$$\begin{aligned} & E\left[\sum_{t=0}^{\infty} \beta^t \tilde{u}\left(x_t, z_t\right)\right] \\ = \; & E\left[\sum_{t=0}^{\infty} \beta^t \left(u\left(0,0\right) + h_x' x_t + h_z' z_t + \frac{1}{2} x_t' H_x x_t + x_t' H_{xz} z_t + \frac{1}{2} z_t' H_z z_t\right)\right], \end{aligned} \tag{86}$$

where h_x is the vector of first derivatives of u with respect to x_t evaluated at the origin, h_z is the vector of first derivatives of u with respect to z_t evaluated at the origin, H_x is the matrix of second derivatives of u with respect to x_t evaluated at the origin, H_z is the matrix of second derivatives of u with respect to z_t evaluated at the origin, and H_{xz} is the matrix of second derivatives of u with respect to x_t and z_t evaluated at the origin. Third, we rewrite equation (86) using condition (25). Let ω_t denote the following vector

$$\omega_t = \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}',$$

and let $\omega_{n,t}$ denote the nth element of ω_t. Condition (25) implies that

$$\sum_{t=0}^{\infty} \beta^t E\left|u\left(0,0\right) + h_x' x_t + h_z' z_t + \frac{1}{2} x_t' H_x x_t + x_t' H_{xz} z_t + \frac{1}{2} z_t' H_z z_t\right| < \infty.$$

It follows that one can change the order of integration and summation on the right-hand side of equation (86):

$$\begin{aligned} & E\left[\sum_{t=0}^{\infty} \beta^t \tilde{u}\left(x_t, z_t\right)\right] \\ = \; & \sum_{t=0}^{\infty} \beta^t E\left[u\left(0,0\right) + h_x' x_t + h_z' z_t + \frac{1}{2} x_t' H_x x_t + x_t' H_{xz} z_t + \frac{1}{2} z_t' H_z z_t\right]. \end{aligned} \tag{87}$$

See Rao (1973), p. 111. Condition (25) also implies that the infinite sum on the right-hand side of equation (87) converges to an element in $\mathbb{R}$. Fourth, we define the vector x_t^*. In each period $t \geq 0$, the vector x_t^* is defined by

$$h_x + H_x x_t^* + H_{xz} z_t = 0. \tag{88}$$

We will show below that H_x is an invertible matrix. Therefore, one can write the last equation as

$$x_t^* = -H_x^{-1} h_x - H_x^{-1} H_{xz} z_t.$$

Hence, x_t^* is uniquely determined and the vector ω_t with $x_t = x_t^*$ satisfies condition (25). Fifth, equation (87) implies that

$$\begin{aligned} & E\left[\sum_{t=0}^{\infty} \beta^t \tilde{u}\left(x_t, z_t\right)\right] - E\left[\sum_{t=0}^{\infty} \beta^t \tilde{u}\left(x_t^*, z_t\right)\right] \\ = & \sum_{t=0}^{\infty} \beta^t E\left[h_x'\left(x_t - x_t^*\right) + \frac{1}{2} x_t' H_x x_t - \frac{1}{2} x_t^{*\prime} H_x x_t^* + \left(x_t - x_t^*\right)' H_{xz} z_t\right]. \end{aligned} \tag{89}$$

Using equation (88) to substitute for $H_{xz} z_t$ in the last equation and rearranging yields

$$\begin{aligned} & E\left[\sum_{t=0}^{\infty} \beta^t \tilde{u}\left(x_t, z_t\right)\right] - E\left[\sum_{t=0}^{\infty} \beta^t \tilde{u}\left(x_t^*, z_t\right)\right] \\ = & \sum_{t=0}^{\infty} \beta^t E\left[\frac{1}{2}\left(x_t - x_t^*\right)' H_x \left(x_t - x_t^*\right)\right]. \end{aligned} \tag{90}$$

Sixth, we compute the vector of first derivatives and the matrices of second derivatives appearing in equations (88) and (90). We obtain

$$h_x = 0, \tag{91}$$

$$H_x = -C^{1-\gamma} \begin{bmatrix} \gamma - 1 + \frac{1}{\alpha}(1+\psi) & 0 & \cdots & \cdots & 0 \\ 0 & 2\frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \cdots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} \\ \vdots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} \\ 0 & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & \cdots & \frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} & 2\frac{1+\Lambda-\alpha}{I(1+\Lambda)\alpha} \end{bmatrix}, \tag{92}$$

and

$$H_{xz} = C^{1-\gamma} \begin{bmatrix} \frac{1}{\alpha}(1+\psi) & 0 \\ 0 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}. \tag{93}$$

Seventh, substituting equations (91)-(93) into equation (88) yields the following system of I equations:

$$c_t^* = \frac{\frac{1}{\alpha}(1+\psi)}{\gamma - 1 + \frac{1}{\alpha}(1+\psi)} a_t, \tag{94}$$

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

$$\hat{c}_{i,t}^* + \sum_{k=1}^{I-1} \hat{c}_{k,t}^* = 0. \tag{95}$$

Finally, we rewrite equation (95). Summing equation (95) over all $i \neq I$ yields

$$\sum_{i=1}^{I-1} \hat{c}_{i,t}^* = 0. \tag{96}$$

Substituting the last equation back into equation (95) yields

$$\hat{c}_{i,t}^* = 0. \tag{97}$$

Collecting equations (90), (92), (94) and (97), we arrive at Proposition 1.

B Proof of Proposition 2

Step 1: We consider rational expectations equilibria with the property that the price level p_t is a linear function of the shocks. In the following, we call this a linear rational expectations equilibrium. When the price level p_t is a linear function of the shocks, the price level can be written as

$$p_t = p_t^a + p_t^\lambda, \tag{98}$$

where p_t^a denotes the component of the price level that is linear in aggregate technology shocks and p_t^λ denotes the component of the price level that is linear in markup shocks. Substituting equations (44), (52) and (98) into equation (47) for the profit-maximizing price of good i in period t yields

$$p_{i,t}^* = (1-\phi_c)\left(p_t^a + p_t^\lambda\right) + \phi_c\left(F_t(L)\varepsilon_t + G_t(L)\nu_t - \frac{\phi_a}{\phi_c}a_t + \frac{\phi_\lambda}{\phi_c}\lambda_t\right). \tag{99}$$

When $F_t(L)\,\varepsilon_t = \frac{\phi_a}{\phi_c} a_t$, equations (45)-(46), (98)-(99) and (48)-(49) imply that in a linear rational expectations equilibrium the component p_t^a is given by

$$p_t^a = (1-\phi_c)\,\frac{1}{I}\sum_{i=1}^{I} E\left[p_t^a|\mathcal{I}_{i,t}\right]. \tag{100}$$

The unique solution to the last equation is $p_t^a = 0$. Hence, when $F_t(L)\,\varepsilon_t = \frac{\phi_a}{\phi_c} a_t$, a linear rational expectations equilibrium has the property that the profit-maximizing price and the price level do not respond to aggregate technology shocks. When the profit-maximizing price does not respond to aggregate technology shocks, price setters in firms put no weight on their signals concerning aggregate technology and thus there is no inefficient price dispersion caused by the noise in the signal concerning aggregate technology. In addition, when $F_t(L)\,\varepsilon_t = \frac{\phi_a}{\phi_c} a_t$ and the price level does not respond to aggregate technology shocks, the equilibrium response of composite consumption to aggregate technology shocks equals the efficient response of composite consumption to aggregate technology shocks. See equations (44), (52) and (43).

Step 2: If $\sigma_\eta^2 > 0$, any monetary policy rule with $F_t(L)\,\varepsilon_t \neq \frac{\phi_a}{\phi_c} a_t$ yields inefficient price dispersion caused by the noise in the signal concerning aggregate technology or an inefficient response of composite consumption to aggregate technology shocks. If price setters in firms put weight on their signals concerning aggregate technology, there is inefficient price dispersion caused by the noise in the signal concerning aggregate technology. If price setters in firms put no weight on their signals concerning aggregate technology, the price level does not respond to aggregate technology shocks. The condition $F_t(L)\,\varepsilon_t \neq \frac{\phi_a}{\phi_c} a_t$ then implies that the equilibrium response of composite consumption to aggregate technology shocks differs from the efficient response of composite consumption to aggregate technology shocks. See equations (44), (52) and (43).

Step 3: The choice of $F_t(L)$ affects neither the equilibrium response of composite consumption to markup shocks nor the extent to which there is inefficient price dispersion caused by the noise in the signal concerning the desired markup.

C Proof of Proposition 3

Step 1: Identical to step 1 in the proof of Proposition 2.

Step 2: If $\mu > 0$, any monetary policy rule with $F_t(L)\,\varepsilon_t \neq \frac{\phi_a}{\phi_c} a_t$ yields inefficient price dispersion caused by the noise in the signal concerning aggregate technology or an inefficient response of composite consumption to aggregate technology shocks. If price setters in firms pay attention to aggregate technology (i.e., $\left(1/\sigma_{\eta}^{2}\right) > 0$), there is inefficient price dispersion caused by the noise in the signal concerning aggregate technology. If price setters in firms pay no attention to aggregate technology (i.e., $\left(1/\sigma_{\eta}^{2}\right) = 0$), the price level does not respond to aggregate technology shocks. The condition $F_t\left(L\right)\varepsilon_t \neq \frac{\phi_a}{\phi_c} a_t$ then implies that the equilibrium response of composite consumption to aggregate technology shocks differs from the efficient response of composite consumption to aggregate technology shocks. See equations (44), (52) and (43).

Step 3: Identical to step 3 in the proof of Proposition 2.

D Proof of Proposition 4

Step 1: Substituting the cash-in-advance constraint (44), $a_t = 0$, the monetary policy $m_t = g_0 \lambda_t$, and $p_t = \theta \lambda_t$ into the equation for the profit-maximizing price (47) yields

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

The price of good i in period t then equals

$$\begin{aligned} p_{i,t} &= \left[\left(1-\phi_c\right)\theta + \phi_c g_0 + \phi_\lambda\right] E\left[\lambda_t | \mathcal{I}_{i,t}\right] \\ &= \left[\left(1-\phi_c\right)\theta + \phi_c g_0 + \phi_\lambda\right] \frac{\sigma_\lambda^2}{\sigma_\lambda^2 + \sigma_\zeta^2}\left(\lambda_t + \zeta_{i,t}\right), \end{aligned} \tag{101}$$

and the price level in period t equals

$$p_t = \left[\left(1-\phi_c\right)\theta + \phi_c g_0 + \phi_\lambda\right] \frac{\sigma_\lambda^2}{\sigma_\lambda^2 + \sigma_\zeta^2}\lambda_t. \tag{102}$$

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

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

Solving the last equation for θ yields

$$\theta = \frac{\phi_c g_0 + \phi_\lambda}{\phi_c + \frac{\sigma_\zeta^2}{\sigma_\lambda^2}}. \tag{103}$$

Substituting equation (103) into equations (101) and (102) yields

$$\begin{aligned} p_t &= \frac{\phi_c g_0 + \phi_\lambda}{\phi_c + \frac{\sigma_\zeta^2}{\sigma_\lambda^2}} \lambda_t, && (104) \\ p_{i,t} - p_t &= \frac{\phi_c g_0 + \phi_\lambda}{\phi_c + \frac{\sigma_\zeta^2}{\sigma_\lambda^2}} \zeta_{i,t}. && (105) \end{aligned}$$

Finally, substituting the monetary policy $m_t = g_0 \lambda_t$ and equation (104) into equation (44) yields

$$c_t = \frac{\frac{\sigma_\zeta^2}{\sigma_\lambda^2} g_0 - \phi_\lambda}{\phi_c + \frac{\sigma_\zeta^2}{\sigma_\lambda^2}} \lambda_t. \tag{106}$$

Step 2: Substituting equations (105), (106), (43) and $a_t = 0$ into the central bank's objective (42) 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

$$g_0^* = \frac{(1 - \delta \phi_c) \phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2} + \delta \phi_c^2}. \tag{107}$$

Step 3: Substituting the optimal monetary policy g_0^* into equations (104) and (106) yields

$$\begin{aligned} p_t &= \frac{\phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2} + \delta \phi_c^2} \lambda_t, && (108) \\ c_t &= -\frac{\delta \phi_c \phi_\lambda}{\frac{\sigma_\zeta^2}{\sigma_\lambda^2} + \delta \phi_c^2} \lambda_t. && (109) \end{aligned}$$

E Proof of Propositions 5 and 6

Step 1: Characterizing equilibrium attention by two equations. We begin by rewriting the equation for the profit-maximizing price (47). Substituting the cash-in-advance constraint (44), $a_t = 0$, the monetary policy $m_t = g_0 \lambda_t$ and $p_t = \theta \lambda_t$ into the equation for the profit-maximizing price (47) yields

$$p_{i,t}^* = \left[(1 - \phi_c) \theta + \phi_c g_0 + \phi_\lambda \right] \lambda_t. \tag{110}$$

Since the profit-maximizing price is given by the last equation and the desired markup follows a white noise process, the attention problem of firm i reads

$$\min_{\kappa\in\mathbb{R}_+}\left\{\frac{\omega}{2}E\left[\left(p_{i,t}-p_{i,t}^*\right)^2\right]+\mu\kappa\right\},$$

subject to

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

$$p_{i,t}=E\left[p_{i,t}^*|s_{\lambda,i,t}\right],$$

$$s_{\lambda,i,t}=\lambda_t+\zeta_{i,t},$$

and

$$\frac{1}{2}\log_2\left(\frac{\sigma_\lambda^2}{\sigma_{\lambda|s_\lambda}^2}\right)=\kappa,$$

where σ_λ^2 denotes the unconditional variance of λ_t and $\sigma_{\lambda|s_\lambda}^2$ denotes the conditional variance of λ_t given $s_{\lambda,i,t}$. Substituting the constraints into the objective, the attention problem of firm i can be expressed as

$$\min_{\kappa\in\mathbb{R}_+}\left\{\frac{\omega}{2}\left[\left(1-\phi_c\right)\theta+\phi_c g_0+\phi_\lambda\right]^2\sigma_\lambda^2 2^{-2\kappa}+\mu\kappa\right\}. \tag{111}$$

The solution to this attention problem is

$$\kappa^*=\begin{cases}\frac{1}{2}\log_2\left(\frac{\omega\left[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda\right]^2\sigma_\lambda^2\ln(2)}{\mu}\right) & \text{if } \frac{\omega\left[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda\right]^2\sigma_\lambda^2\ln(2)}{\mu}\geq 1\\ 0 & \text{otherwise}\end{cases}. \tag{112}$$

The price set by firm i in period t then equals

$$\begin{aligned}p_{i,t} &= \left[\left(1-\phi_c\right)\theta+\phi_c g_0+\phi_\lambda\right]E\left[\lambda_t|s_{\lambda,i,t}\right]\\ &= \left[\left(1-\phi_c\right)\theta+\phi_c g_0+\phi_\lambda\right]\frac{\frac{\sigma_\lambda^2}{\sigma_\zeta^2}}{\frac{\sigma_\lambda^2}{\sigma_\zeta^2}+1}\left(\lambda_t+\zeta_{i,t}\right),\end{aligned} \tag{113}$$

where

$$\frac{\sigma_\lambda^2}{\sigma_\zeta^2}=2^{2\kappa^*}-1. \tag{114}$$

The price level in period t equals

$$p_t=\left[\left(1-\phi_c\right)\theta+\phi_c g_0+\phi_\lambda\right]\left(1-2^{-2\kappa^*}\right)\lambda_t. \tag{115}$$

Thus, the set of rational expectations equilibria of the form $p_t = \theta\lambda_t$ is given by the solutions to the following two equations:

$$\theta = \left[\left(1-\phi_c\right)\theta + \phi_c g_0 + \phi_\lambda\right]\left(1-2^{-2\kappa^*}\right), \tag{116}$$

and

$$\kappa^* = \begin{cases} \frac{1}{2}\log_2\left(\frac{\omega[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]^2\sigma_\lambda^2\ln(2)}{\mu}\right) & \text{if } \frac{\omega[(1-\phi_c)\theta+\phi_c g_0+\phi_\lambda]^2\sigma_\lambda^2\ln(2)}{\mu} \geq 1 \\ 0 & \text{otherwise} \end{cases}. \tag{117}$$

Equation (116) determines θ (the responsiveness of the price level to the desired markup) as a function of κ^* (equilibrium attention), while equation (117) determines κ^* as a function of θ. Solving equation (116) for θ yields

$$\theta = \frac{\left(\phi_c g_0 + \phi_\lambda\right)\left(1-2^{-2\kappa^*}\right)}{1-\left(1-\phi_c\right)\left(1-2^{-2\kappa^*}\right)}. \tag{118}$$

The set of rational expectations equilibria of the form $p_t = \theta\lambda_t$ for given monetary policy g_0 consists of the pairs (κ^*, θ) that solve equations (117)-(118).

Step 2: **Zero attention equilibrium**. We now study under which conditions there exists a solution to equations (117)-(118) with the property $\kappa^* = 0$. We call this a zero attention equilibrium. It follows from equation (118) that $\kappa^* = 0$ implies $\theta = 0$. Furthermore, it follows from equation (117) that at $\theta = 0$ we have $\kappa^* = 0$ if and only if

$$\frac{\omega\left(\phi_c g_0 + \phi_\lambda\right)^2\sigma_\lambda^2\ln(2)}{\mu} \leq 1. \tag{119}$$

Hence, there exists a rational expectations equilibrium of the form $p_t = \theta\lambda_t$ with $\kappa^* = 0$ if and only if condition (119) is satisfied. Note that the central bank can always ensure the existence of a zero attention equilibrium by making the term $\left(\phi_c g_0 + \phi_\lambda\right)^2$ sufficiently small through an appropriate choice of g_0.

Step 3: Interior attention equilibrium. Next we study under which conditions there exists a solution to equations (117)-(118) with the property

$$\kappa^* = \frac{1}{2}\log_2\left(\frac{\omega\left[\left(1-\phi_c\right)\theta + \phi_c g_0 + \phi_\lambda\right]^2\sigma_\lambda^2\ln(2)}{\mu}\right). \tag{120}$$

We call this an interior attention equilibrium because in such an equilibrium the non-negativity constraint $\kappa \geq 0$ in the firms' attention problem (111) is not binding. Substituting equation (118) into equation (120) yields

$$\kappa^* = \frac{1}{2}\log_2\left(\frac{\omega\frac{(\phi_c g_0+\phi_\lambda)^2}{\left[1-(1-\phi_c)\left(1-2^{-2\kappa^*}\right)\right]^2}\sigma_\lambda^2\ln(2)}{\mu}\right). \tag{121}$$

Rearranging the last equation yields a quadratic equation in 2^{κ^*}:

$$\phi_c\left(2^{\kappa^*}\right)^2 - \sqrt{\frac{\omega\left(\phi_c g_0+\phi_\lambda\right)^2\sigma_\lambda^2\ln(2)}{\mu}}2^{\kappa^*} + 1 - \phi_c = 0.$$

Defining $x \equiv 2^{\kappa^*}$, the last equation can be written as

$$\phi_c x^2 - \sqrt{\frac{\omega\left(\phi_c g_0+\phi_\lambda\right)^2\sigma_\lambda^2\ln(2)}{\mu}}x + 1 - \phi_c = 0. \tag{122}$$

An interior attention equilibrium has to satisfy this quadratic equation as well as: $x \in \mathbb{R}$ and $x \geq 1$. Define

$$b \equiv \sqrt{\frac{\omega\left(\phi_c g_0+\phi_\lambda\right)^2\sigma_\lambda^2\ln(2)}{\mu}}. \tag{123}$$

The quadratic equation (122) has two solutions:

$$x_H = \frac{b+\sqrt{b^2-4\phi_c(1-\phi_c)}}{2\phi_c}, \tag{124}$$

and

$$x_L = \frac{b-\sqrt{b^2-4\phi_c(1-\phi_c)}}{2\phi_c}. \tag{125}$$

We now check whether these two solutions to the quadratic equation (122) satisfy: $x \in \mathbb{R}$ and $x \geq 1$. First, consider the case of $\phi_c \in \left(0, \frac{1}{2}\right]$. Then x_H and x_L are real if and only if $b \geq \sqrt{4\phi_c(1-\phi_c)}$. At $b = \sqrt{4\phi_c(1-\phi_c)}$, we have $x_H = x_L = \sqrt{\frac{1}{\phi_c}-1} \geq 1$. Furthermore, x_H is increasing in b and thus $x_H \geq 1$ for all $b \geq \sqrt{4\phi_c(1-\phi_c)}$, whereas x_L is decreasing in b and $x_L \geq 1$ for all $b \in \left[\sqrt{4\phi_c(1-\phi_c)}, 1\right]$. Hence, if $\phi_c \in \left(0, \frac{1}{2}\right]$, then x_H is an interior attention equilibrium so long as $b \geq \sqrt{4\phi_c(1-\phi_c)}$, while x_L is an interior attention equilibrium so long as $b \in \left[\sqrt{4\phi_c(1-\phi_c)}, 1\right]$. Second, consider the case of $\phi_c \in \left(\frac{1}{2}, 1\right]$. Again x_H and x_L are real if and only if $b \geq \sqrt{4\phi_c(1-\phi_c)}$. At $b = \sqrt{4\phi_c(1-\phi_c)}$, we have $x_H = x_L = \sqrt{\frac{1}{\phi_c}-1} < 1$. Furthermore, x_H is increasing in b and $x_H \geq 1$ for all $b \geq 1$, whereas x_L is non-increasing in b and thus $x_L < 1$ for all $b \geq \sqrt{4\phi_c(1-\phi_c)}$. Hence, if $\phi_c \in \left(\frac{1}{2}, 1\right]$, then x_H is an interior attention equilibrium so long as $b \geq 1$, while x_L is not an interior attention equilibrium. Finally, consider the case of $\phi_c > 1$. Then x_H and x_L are real for all $b \geq 0$. At $b = 0$, we have $x_H = \sqrt{1 - \frac{1}{\phi_c}} < 1$ and $x_L = -\sqrt{1 - \frac{1}{\phi_c}} < 0$. Furthermore, x_H is increasing in b and $x_H \geq 1$ for all $b \geq 1$, whereas $x_L < 0$ for all $b \geq 0$. Hence, if $\phi_c > 1$, then x_H is an interior attention equilibrium so long as $b \geq 1$, while x_L is not an interior attention equilibrium. In summary, if and only if either $\phi_c \in \left(0, \frac{1}{2}\right]$ and $b \geq \sqrt{4\phi_c (1 - \phi_c)}$ or $\phi_c > \frac{1}{2}$ and $b \geq 1$, then x_H is an interior attention equilibrium. In addition, if and only if $\phi_c \in \left(0, \frac{1}{2}\right]$ and $b \in \left[\sqrt{4\phi_c (1 - \phi_c)}, 1\right]$, then x_L is an interior attention equilibrium.

Step 4: Uniqueness and multiplicity of equilibria. When $\phi_c \geq \frac{1}{2}$, there exists a unique rational expectations equilibrium of the form $p_t = \theta\lambda_t$ for any monetary policy $g_0 \in \mathbb{R}$. In particular, if $b \in [0, 1)$ then $\kappa^* = 0$ is the unique equilibrium; if $b = 1$ then $\kappa^* = \log_2 (x_H) = 0$ is the unique equilibrium; and if $b > 1$ then $\kappa^* = \log_2 (x_H)$ is the unique equilibrium. By contrast, when $\phi_c \in \left(0, \frac{1}{2}\right)$, there exist multiple rational expectations equilibria of the form $p_t = \theta\lambda_t$ for some monetary policies $g_0 \in \mathbb{R}$. In particular, if $b \in \left[0, \sqrt{4\phi_c (1 - \phi_c)}\right)$ then $\kappa^* = 0$ is the unique equilibrium; if $b = \sqrt{4\phi_c (1 - \phi_c)}$ then $\kappa^* = 0$ and $\kappa^* = \log_2 (x_L) = \log_2 (x_H) = \log_2 \left(\sqrt{\frac{1}{\phi_c} - 1}\right)$ are equilibria; if $b \in \left(\sqrt{4\phi_c (1 - \phi_c)}, 1\right)$ then $\kappa^* = 0$, $\kappa^* = \log_2 (x_L)$ and $\kappa^* = \log_2 (x_H)$ are equilibria where x_L is decreasing in b and x_H is increasing in b; if $b = 1$ then $\kappa^* = \log_2 (x_L) = 0$ and $\kappa^* = \log_2 (x_H) = \log_2 \left(\frac{1}{\phi_c} - 1\right)$ are equilibria; and if $b > 1$ then $\kappa^* = \log_2 (x_H)$ is the unique equilibrium. See steps 2 and 3.

Step 5: Price dispersion and consumption variance. We now derive expressions for price dispersion and consumption variance at an equilibrium. First, we derive expressions for individual prices and the price level. Substituting equations (114) and (118) into equation (113) yields

$$p_{i,t} = \frac{(\phi_c g_0 + \phi_\lambda)\left(1 - 2^{-2\kappa^*}\right)}{1 - (1 - \phi_c)\left(1 - 2^{-2\kappa^*}\right)} \left(\lambda_t + \zeta_{i,t}\right). \tag{126}$$

Substituting equation (118) into equation (115) yields

$$p_t = \frac{(\phi_c g_0 + \phi_\lambda)\left(1 - 2^{-2\kappa^*}\right)}{1 - (1 - \phi_c)\left(1 - 2^{-2\kappa^*}\right)} \lambda_t. \tag{127}$$

Second, we derive a simple expression for price dispersion at an equilibrium. Consider the case of an equilibrium with $\kappa^* > 0$. An equilibrium with $\kappa^* > 0$ is an interior attention equilibrium and in an interior attention equilibrium equation (121) holds. Equations (126) and (127) imply

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

Substituting equation (121) into the last equation yields

$$E\left[(p_{i,t}-p_t)^2\right]=\frac{\frac{\mu}{\omega}}{\sigma_\lambda^2\ln(2)}2^{2\kappa^*}\left(1-2^{-2\kappa^*}\right)^2\sigma_\zeta^2.$$

Furthermore, substituting equation (114) into the last equation and rearranging yields

$$E\left[(p_{i,t}-p_t)^2\right]=\frac{\frac{\mu}{\omega}}{\ln(2)}\left(1-2^{-2\kappa^*}\right). \tag{128}$$

Next consider the case of an equilibrium with $\kappa^*=0$. Equation (128) holds again because in an equilibrium with $\kappa^*=0$ we have $E\left[(p_{i,t}-p_t)^2\right]=0$. In summary, in any equilibrium, price dispersion is given by equation (128). It follows that equilibrium price dispersion is an increasing function of equilibrium attention. Third, we derive an expression for consumption variance at an equilibrium. Substituting the monetary policy $m_t=g_0\lambda_t$ and the equation for the price level (127) into the cash-in-advance constraint (44) yields

$$c_t=\left[g_0-\frac{(\phi_c g_0+\phi_\lambda)\left(1-2^{-2\kappa^*}\right)}{1-(1-\phi_c)\left(1-2^{-2\kappa^*}\right)}\right]\lambda_t, \tag{129}$$

implying

$$E\left[c_t^2\right]=\left[g_0-\frac{(\phi_c g_0+\phi_\lambda)\left(1-2^{-2\kappa^*}\right)}{1-(1-\phi_c)\left(1-2^{-2\kappa^*}\right)}\right]^2\sigma_\lambda^2. \tag{130}$$

The first term in square brackets in equation (129) equals the response of nominal spending to the desired markup, while the second term in square brackets in equation (129) equals the response of the price level to the desired markup. The difference between the two determines the response of composite consumption to the desired markup.

Step 6: Optimal monetary policy has to satisfy $g_0\geq-\frac{\phi_\lambda}{\phi_c}$. Having characterized the set of rational expectations equilibria of the form $p_t=\theta\lambda_t$ for given monetary policy g_0 and having derived expressions for price dispersion and consumption variance, we now derive results concerning optimal monetary policy. We begin by showing that optimal monetary policy has to satisfy $g_0\geq-\frac{\phi_\lambda}{\phi_c}$. The proof is as follows. First, at the monetary policy $g_0=-\frac{\phi_\lambda}{\phi_c}$ we have $b=0$ and thus the unique rational expectations equilibrium of the form $p_t=\theta\lambda_t$ is a zero attention equilibrium, implying that price dispersion equals zero and consumption variance equals $E\left[c_t^2\right]=\left(\frac{\phi_\lambda}{\phi_c}\right)^2\sigma_\lambda^2$. Second, consider a monetary policy $g_0<-\frac{\phi_\lambda}{\phi_c}$. Price dispersion at a monetary policy $g_0<-\frac{\phi_\lambda}{\phi_c}$ is weakly larger than price dispersion at the monetary policy $g_0=-\frac{\phi_\lambda}{\phi_c}$ because price dispersion is always weakly larger than zero. Furthermore, consumption variance at a monetary policy $g_0<-\frac{\phi_\lambda}{\phi_c}$ is strictly larger than consumption variance at the monetary policy $g_0 = -\frac{\phi_\lambda}{\phi_c}$. This result follows from the fact that consumption variance at an equilibrium is given by equation (130) and, for all $g_0 < -\frac{\phi_\lambda}{\phi_c}$, we have

$$g_0 - \frac{(\phi_c g_0 + \phi_\lambda)\left(1 - 2^{-2\kappa^*}\right)}{1 - (1 - \phi_c)\left(1 - 2^{-2\kappa^*}\right)} < -\frac{\phi_\lambda}{\phi_c} < 0. \tag{131}$$

In summary, a monetary policy $g_0 < -\frac{\phi_\lambda}{\phi_c}$ yields weakly larger price dispersion and strictly larger consumption variance than the monetary policy $g_0 = -\frac{\phi_\lambda}{\phi_c}$. Hence, a monetary policy $g_0 < -\frac{\phi_\lambda}{\phi_c}$ cannot be optimal. To see what this result means economically, note that the condition $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$ can be written as $\phi_c g_0 + \phi_\lambda \geq 0$. Furthermore, substituting equation (118) into equation (110) yields the following equation for the profit-maximizing price at an equilibrium

$$p_{i,t}^* = \frac{\phi_c g_0 + \phi_\lambda}{1 - (1 - \phi_c)\left(1 - 2^{-2\kappa^*}\right)} \lambda_t. \tag{132}$$

The result that optimal monetary policy has to satisfy $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$ means that at an optimal monetary policy the profit-maximizing price cannot be decreasing in the desired markup, implying that individual prices and the price level cannot be decreasing in the desired markup.

Step 7: Optimal monetary policy when $\phi_c \geq \frac{1}{2}$. First, when $\phi_c \geq \frac{1}{2}$, there exists a unique rational expectations equilibrium of the form $p_t = \theta \lambda_t$ for any monetary policy $g_0 \in \mathbb{R}$: if $b \in [0, 1)$ then $\kappa^* = 0$ is the unique equilibrium; if $b = 1$ then $\kappa^* = \log_2(x_H) = 0$ is the unique equilibrium; and if $b > 1$ then $\kappa^* = \log_2(x_H)$ is the unique equilibrium. See step 4. Second, in the derivation of optimal monetary policy we can focus on $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$. See step 6. Furthermore, equation (123) implies that the variable b is an increasing function of the monetary policy g_0 for all $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$. Define $\bar{g}_0$ as the value of $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \infty\right)$ at which $b = 1$. Equation (123) implies that

$$\bar{g}_0 = -\frac{\phi_\lambda}{\phi_c} + \frac{1}{\phi_c}\sqrt{\frac{\mu}{\omega \sigma_\lambda^2 \ln(2)}}. \tag{133}$$

If $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \bar{g}_0\right)$ then $\kappa^* = 0$ is the unique equilibrium; if $g_0 = \bar{g}_0$ then $\kappa^* = \log_2(x_H) = 0$ is the unique equilibrium; and if $g_0 > \bar{g}_0$ then $\kappa^* = \log_2(x_H)$ is the unique equilibrium. Note that the condition $\frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} \leq 1$ implies $\bar{g}_0 \geq 0$, whereas the condition $\frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} > 1$ implies $\bar{g}_0 < 0$. Hence, when $\frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} \leq 1$ the monetary policy $g_0 = 0$ yields a zero attention equilibrium, whereas when $\frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} > 1$ the central bank has to lower the money supply after a positive markup shock to attain a zero attention equilibrium. Third, consider the case of $\phi_c \geq \frac{1}{2}$ and $\frac{\omega \phi_\lambda^2 \sigma_\lambda^2 \ln(2)}{\mu} \leq 1$. At the monetary policy $g_0 = 0$, we have $b \leq 1$ and therefore $\kappa^* = 0$ is the unique equilibrium, implying that price dispersion equals zero, $E\left[(p_{i,t}-p_t)^2\right]=0$, and consumption variance equals zero, $E\left[c_t^2\right]=g_0^2\sigma_\lambda^2=0$. See equations (128) and (130). Thus, in the case of $\phi_c\geq\frac{1}{2}$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu}\leq 1$ the monetary policy $g_0=0$ attains the efficient allocation. Furthermore, any monetary policy $g_0\neq 0$ does not attain the efficient allocation. If the equilibrium at the monetary policy $g_0\neq 0$ is an equilibrium with $\kappa^*=0$ then consumption variance is strictly positive, while if the equilibrium at the monetary policy $g_0\neq 0$ is an equilibrium with $\kappa^*>0$ then price dispersion is strictly positive. See equations (128) and (130). Hence, in the case of $\phi_c\geq\frac{1}{2}$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu}\leq 1$ the unique optimal monetary policy is $g_0^*=0$. Fourth, consider the case of $\phi_c\geq\frac{1}{2}$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu}>1$. We derive optimal monetary policy in this case by showing that the monetary policy minimizing objective (42) among all monetary policies $g_0\in\left[-\frac{\phi_\lambda}{\phi_c},\bar{g}_0\right]$ is $g_0=\bar{g}_0$ and by showing that the monetary policy minimizing objective (42) among all monetary policies $g_0\in[\bar{g}_0,\infty)$ is $g_0=\bar{g}_0$. Combining results then yields that in the case of $\phi_c\geq\frac{1}{2}$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu}>1$ the unique optimal monetary policy among all monetary policies $g_0\in\mathbb{R}$ is $g_0=\bar{g}_0$. For all $g_0\in\left[-\frac{\phi_\lambda}{\phi_c},\bar{g}_0\right]$, we have $b\leq 1$ and thus $\kappa^*=0$ is the unique equilibrium, implying that price dispersion equals zero and consumption variance equals $E\left[c_t^2\right]=g_0^2\sigma_\lambda^2$. Furthermore, $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu}>1$ implies $\bar{g}_0<0$. It follows that in the case of $\phi_c\geq\frac{1}{2}$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu}>1$ the monetary policy minimizing objective (42) among all monetary policies $g_0\in\left[-\frac{\phi_\lambda}{\phi_c},\bar{g}_0\right]$ is $g_0=\bar{g}_0$. Next, for all $g_0\in[\bar{g}_0,\infty)$, we have $b\geq 1$ and thus $\kappa^*=\log_2(x_H)$ is the unique equilibrium. Let us study price dispersion. Since equilibrium price dispersion is strictly increasing in κ^*, $\kappa^*=\log_2(x_H)$, x_H is strictly increasing in b, and b is strictly increasing in g_0 for all $g_0\geq\bar{g}_0$, it follows that price dispersion is strictly increasing in g_0 for all $g_0\geq\bar{g}_0$. See equations (123), (124) and (128). Let us turn to consumption variance. Equilibrium consumption is given by equation (129). Furthermore, when $\kappa^*=\log_2(x_H)$, equation (121) holds. Rearranging equation (121) using $g_0\geq-\frac{\phi_\lambda}{\phi_c}$ yields

$$\frac{\phi_c g_0+\phi_\lambda}{1-(1-\phi_c)\left(1-2^{-2\kappa^*}\right)}=\sqrt{\frac{\mu}{\omega\sigma_\lambda^2\ln(2)}}2^{\kappa^*}. \tag{134}$$

Substituting the last equation into equation (129) yields

$$c_t=\left[g_0-\sqrt{\frac{\mu}{\omega\sigma_\lambda^2\ln(2)}}\left(2^{\kappa^*}-2^{-\kappa^*}\right)\right]\lambda_t. \tag{135}$$

In addition, solving the definition of the variable b (i.e., equation (123)) for g_0 using $g_0\geq-\frac{\phi_\lambda}{\phi_c}$

yields

$$g_0 = -\frac{\phi_\lambda}{\phi_c} + \frac{b}{\phi_c}\sqrt{\frac{\mu}{\omega\sigma_\lambda^2 \ln(2)}}. \tag{136}$$

Substituting equation (136), $\kappa^* = \log_2(x_H)$ and equation (124) into equation (135) yields

$$c_t = \left[-\frac{\phi_\lambda}{\phi_c} + \sqrt{\frac{\mu}{\omega\sigma_\lambda^2 \ln(2)}} \left(\frac{b}{\phi_c} - \frac{b + \sqrt{b^2 - 4\phi_c(1-\phi_c)}}{2\phi_c} + \frac{1}{\frac{b+\sqrt{b^2-4\phi_c(1-\phi_c)}}{2\phi_c}} \right) \right] \lambda_t.$$

Rearranging the last equation yields

$$c_t = \left[-\frac{\phi_\lambda}{\phi_c} + \sqrt{\frac{\mu}{\omega\sigma_\lambda^2 \ln(2)}} \frac{2}{b + \sqrt{b^2 - 4\phi_c(1-\phi_c)}} \right] \lambda_t. \tag{137}$$

Hence, when $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$ and $\kappa^* = \log_2(x_H)$, equilibrium consumption is given by equation (137). The term in square brackets in equation (137) is strictly decreasing in b for all $b \geq 1$. Furthermore, in the case of $\phi_c \geq \frac{1}{2}$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu} > 1$, the term in square brackets in equation (137) is strictly negative at $b = 1$. Thus, in the case of $\phi_c \geq \frac{1}{2}$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu} > 1$, consumption variance is strictly increasing in g_0 for all $g_0 \geq \bar{g}_0$. It follows that, in the case of $\phi_c \geq \frac{1}{2}$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu} > 1$, the monetary policy minimizing objective (42) among all monetary policies $g_0 \in [\bar{g}_0, \infty)$ is $g_0 = \bar{g}_0$ because both price dispersion and consumption variance are strictly increasing in g_0 for all $g_0 \geq \bar{g}_0$. Combining results yields that in the case of $\phi_c \geq \frac{1}{2}$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2\ln(2)}{\mu} > 1$ the unique optimal monetary policy among all monetary policies $g_0 \in \mathbb{R}$ is $g_0 = \bar{g}_0$.

Step 8: Optimal monetary policy when $\phi_c \in \left(0, \frac{1}{2}\right)$. First, when $\phi_c \in \left(0, \frac{1}{2}\right)$, there exist multiple rational expectations equilibria of the form $p_t = \theta\lambda_t$ for some monetary policies $g_0 \in \mathbb{R}$. We will use the following results below. If $b \in \left[0, \sqrt{4\phi_c(1-\phi_c)}\right]$ then $\kappa^* = 0$ is an equilibrium. If $b \geq 1$ then $\kappa^* = \log_2(x_H)$ is an equilibrium. Furthermore, if $b \in \left[0, \sqrt{4\phi_c(1-\phi_c)}\right)$ or $b > 1$ then there exists a unique equilibrium, while if $b \in \left[\sqrt{4\phi_c(1-\phi_c)}, 1\right]$ then there exist multiple equilibria. See step 4. Second, in the derivation of optimal monetary policy we can focus on $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$. See step 6. Furthermore, equation (123) implies that b is strictly increasing in g_0 for all $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$. Define $\hat{g}_0$ as the value of $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \infty\right)$ at which $b = \sqrt{4\phi_c(1-\phi_c)}$. Define $\bar{g}_0$ as the value of $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \infty\right)$ at which $b = 1$. Formally,

$$\hat{g}_0 = -\frac{\phi_\lambda}{\phi_c} + \frac{\sqrt{4\phi_c(1-\phi_c)}}{\phi_c}\sqrt{\frac{\mu}{\omega\sigma_\lambda^2 \ln(2)}}, \tag{138}$$

and

$$\bar{g}_0 = -\frac{\phi_\lambda}{\phi_c} + \frac{1}{\phi_c}\sqrt{\frac{\mu}{\omega\sigma_\lambda^2 \ln(2)}}. \tag{139}$$

For all $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \hat{g}_0\right)$, $\kappa^* = 0$ is the unique equilibrium. For all $g_0 > \bar{g}_0$, $\kappa^* = \log_2(x_H)$ is the unique equilibrium. Note that the condition $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} < 4\phi_c(1-\phi_c)$ implies $\hat{g}_0 > 0$. Third, consider the case of $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} < 4\phi_c(1-\phi_c)$. At the monetary policy $g_0 = 0$, we have $g_0 < \hat{g}_0$ and thus $\kappa^* = 0$ is the unique equilibrium, implying that price dispersion equals zero and consumption variance equals zero, $E\left[c_t^2\right] = g_0^2\sigma_\lambda^2 = 0$. See equations (128) and (130). Thus, when $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} < 4\phi_c(1-\phi_c)$ the monetary policy $g_0 = 0$ attains the efficient allocation as the unique equilibrium allocation. Moreover, any monetary policy $g_0 \neq 0$ does not attain the efficient allocation. If the equilibrium at the monetary policy $g_0 \neq 0$ is an equilibrium with $\kappa^* = 0$ then consumption variance is strictly positive, while if the equilibrium at the monetary policy $g_0 \neq 0$ is an equilibrium with $\kappa^* > 0$ then price dispersion is strictly positive. See equations (128) and (130). Hence, when $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} < 4\phi_c(1-\phi_c)$ the unique optimal monetary policy is $g_0^* = 0$. Fourth, consider the case of $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$. For all monetary policies $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \hat{g}_0\right]$, $\kappa^* = 0$ is an equilibrium. Let us rank these zero attention equilibria for different monetary policies $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \hat{g}_0\right]$. In a zero attention equilibrium price dispersion equals zero and consumption variance equals $E\left[c_t^2\right] = g_0^2\sigma_\lambda^2$. The condition $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$ implies $\hat{g}_0 \leq 0$. Hence, in the case of $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$, the value of objective (42) at an equilibrium with $\kappa^* = 0$ is strictly decreasing and continuous in g_0 for all $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \hat{g}_0\right]$. Next, for all monetary policies $g_0 \geq \bar{g}_0$, $\kappa^* = \log_2(x_H)$ is an equilibrium. Let us rank these equilibria with $\kappa^* = \log_2(x_H)$ for different monetary policies $g_0 \geq \bar{g}_0$. It follows from equations (128), (123) and (124) that price dispersion at an equilibrium with $\kappa^* = \log_2(x_H)$ is strictly increasing in g_0 for all $g_0 \geq \bar{g}_0$. Furthermore, the same derivation as in step 7 yields that when $g_0 \geq -\frac{\phi_\lambda}{\phi_c}$ and $\kappa^* = \log_2(x_H)$ equilibrium consumption equals

$$c_t = \left[-\frac{\phi_\lambda}{\phi_c} + \sqrt{\frac{\mu}{\omega\sigma_\lambda^2 \ln(2)}}\frac{2}{b + \sqrt{b^2 - 4\phi_c(1-\phi_c)}}\right]\lambda_t. \tag{140}$$

The term in square brackets in equation (140) is strictly decreasing in b for all $b \geq 1$. Moreover, in the case of $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$, the term in square brackets in equation (140) is strictly negative at $b = 1$. Hence, in the case of $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$, both price dispersion and consumption variance at an equilibrium with $\kappa^* = \log_2(x_H)$ are strictly increasing in g_0 for all $g_0 \geq \bar{g}_0$, implying that the value of objective (42) at an equilibrium with $\kappa^* = \log_2(x_H)$ is strictly increasing in g_0 for all $g_0 \geq \bar{g}_0$. Finally, in the case of $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$, we compare the equilibrium with $\kappa^* = 0$ at $g_0 = \hat{g}_0$ to the equilibrium with $\kappa^* = \log_2(x_H)$ at $g_0 = \bar{g}_0$. At $g_0 = \bar{g}_0$ we have $b = 1$ and thus $\log_2(x_H) = \log_2\left(\frac{1}{\phi_c} - 1\right) > 0$ in the case of $\phi_c \in \left(0, \frac{1}{2}\right)$. See equation (124). It follows from equation (128) that price dispersion is strictly smaller in a zero attention equilibrium than in the equilibrium with $\kappa^* = \log_2(x_H)$ at $g_0 = \bar{g}_0$. In addition, consumption variance is strictly smaller in the zero attention equilibrium at $g_0 = \hat{g}_0$ than in the equilibrium with $\kappa^* = \log_2(x_H)$ at $g_0 = \bar{g}_0$ because the conditions $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$ imply

$$\left[-\frac{\phi_\lambda}{\phi_c} + \frac{\sqrt{4\phi_c(1-\phi_c)}}{\phi_c}\sqrt{\frac{\mu}{\omega\sigma_\lambda^2 \ln(2)}}\right]^2 \sigma_\lambda^2 < \left[-\frac{\phi_\lambda}{\phi_c} + \sqrt{\frac{\mu}{\omega\sigma_\lambda^2 \ln(2)}}\frac{2}{1+\sqrt{(1-2\phi_c)^2}}\right]^2 \sigma_\lambda^2.$$

Hence, in the case of $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$, the zero attention equilibrium at $g_0 = \hat{g}_0$ yields a strictly smaller value of objective (42) than the equilibrium with $\kappa^* = \log_2(x_H)$ at $g_0 = \bar{g}_0$. In summary, in the case of $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$, we have the following four results: (i) there exists a unique rational expectations equilibrium of the form $p_t = \theta\lambda_t$ for all $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \hat{g}_0\right)$ and $g_0 > \bar{g}_0$, whereas there exist multiple rational expectations equilibria of the form $p_t = \theta\lambda_t$ for all $g_0 \in [\hat{g}_0, \bar{g}_0]$, (ii) the value of objective (42) at an equilibrium with $\kappa^* = 0$ is strictly decreasing and continuous in g_0 for all $g_0 \in \left[-\frac{\phi_\lambda}{\phi_c}, \hat{g}_0\right]$, (iii) the value of objective (42) at an equilibrium with $\kappa^* = \log_2(x_H)$ is strictly increasing in g_0 for all $g_0 \geq \bar{g}_0$, and (iv) the equilibrium with $\kappa^* = 0$ at $g_0 = \hat{g}_0$ yields a strictly smaller value of objective (42) than the equilibrium with $\kappa^* = \log_2(x_H)$ at $g_0 = \bar{g}_0$. It follows that, in the case of $\phi_c \in \left(0, \frac{1}{2}\right)$ and $\frac{\omega\phi_\lambda^2\sigma_\lambda^2 \ln(2)}{\mu} \geq 4\phi_c(1-\phi_c)$, the best the central bank can do among all monetary policies $g_0 \in \mathbb{R}$ if the central bank wants to obtain a unique equilibrium of the form $p_t = \theta\lambda_t$ is to choose a g_0 marginally below $\hat{g}_0$. At this policy, price setters in firms devote no attention to the desired markup, the price level does not respond to markup shocks, and there is no inefficient price dispersion.

References

[1] Adam, Klaus (2007). "Optimal Monetary Policy with Imperfect Common Knowledge," *Journal of Monetary Economics*, 54(2), 267-301.

[2] Angeletos, George-Marios, and Jennifer La'O (2008). "Dispersed Information over the Business Cycle: Optimal Fiscal and Monetary Policy," Discussion paper, MIT.

[3] Angeletos, George-Marios, and Alessandro Pavan (2007). "Efficient Use of Information and Social Value of Information," *Econometrica*, 75(4), 1103-1142.

[4] Ball, Laurence, N. Gregory Mankiw, and Ricardo Reis (2005). "Monetary Policy for Inattentive Economies," *Journal of Monetary Economics*, 52(4), 703-725.

[5] Calvo, Guillermo (1983). "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12, 383-398.

[6] Dixit, Avinash, and Joseph E. Stiglitz (1977). "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67(3), 297-308.

[7] Gali, Jordi (2008). Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework. Princeton: Princeton University Press.

[8] Giannoni, Marc, and Michael Woodford (2010). "Optimal Target Criteria for Stabilization Policy," Discussion paper, Columbia University.

[9] Golosov, Mikhail, and Robert E. Lucas, Jr. (2007). "Menu Costs and Phillips Curves," Journal of Political Economy, 115(2), 171-199.

[10] Hellwig, Christian (2005). "Heterogeneous Information and the Welfare Effects of Public Information Disclosures," Discussion paper, UCLA.

[11] Hellwig, Christian, and Laura Veldkamp (2009). "Knowing What Others Know: Coordination Motives in Information Acquisition," *Review of Economic Studies*, 76(1), 223-251.

[12] Lorenzoni, Guido (2010). "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information," *Review of Economic Studies*, 77(1), 305-338.

[13] Luo, Yulei (2008). "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics*, 11(2), 366-385.

[14] Maćkowiak, Bartosz, and Mirko Wiederholt (2009). "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99(3), 769-803.

[15] Maćkowiak, Bartosz, and Mirko Wiederholt (2010). "Business Cycle Dynamics under Rational Inattention," Discussion paper, ECB and Northwestern University.

[16] Mankiw, N. Gregory, and Ricardo Reis (2002). "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117(4), 1295-1328.

[17] Matejka, Filip (2010). "Rationally Inattentive Seller: Sales and Discrete Pricing," Discussion paper, Princeton University.

[18] Midrigan, Virgiliu (2010). "Menu Costs, Multi-Product Firms, and Aggregate Fluctuations," *Econometrica*, forthcoming.

[19] Mondria, Jordi (2010). "Portfolio Choice, Attention Allocation, and Price Comovement," *Journal of Economic Theory*, forthcoming.

[20] Morris, Stephen, and Hyun Song Shin (2002). "Social Value of Public Information," *American Economic Review*, 92(5), 1521-1534.

[21] Nakamura, Emi, and Jón Steinsson (2008). "Five Facts about Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics*, 123(4), 1415-1464.

[22] Paciello, Luigi (2010). "Monetary Policy Activism and Price Responsiveness to Aggregate Shocks under Rational Inattention," Discussion paper, Einaudi Institute for Economics and Finance.

[23] Rao, C. R. (1973). Linear Statistical Inference and its Applications, New York: Wiley.

[24] Sims, Christopher A. (2003). "Implications of Rational Inattention," *Journal of Monetary Economics*, 50(3), 665-690.

[25] Sims, Christopher A. (2006). "Rational Inattention: Beyond the Linear-Quadratic Case," *American Economic Review Papers and Proceedings*, 96(2), 158-163.

[26] Sims, Christopher A. (2010). "Rational Inattention and Monetary Economics," Prepared for Handbook of Monetary Economics, Elsevier.

[27] Svensson, Lars E. O., and Michael Woodford (2003). "Indicator Variables for Optimal Policy," *Journal of Monetary Economics*, 50(3), 691-720.

[28] Svensson, Lars E. O., and Michael Woodford (2004). "Indicator Variables for Optimal Policy under Asymmetric Information," *Journal of Economic Dynamics and Control*, 28(4), 661-690.

[29] Van Nieuwerburgh, Stijn, and Laura Veldkamp (2009). "Information Immobility and the Home Bias Puzzle," *Journal of Finance*, 64(3), 1187-1215.

[30] Van Nieuwerburgh, Stijn, and Laura Veldkamp (2010). "Information Acquisition and Under-Diversification," *Review of Economic Studies*, 77(2), 779-805.

[31] Woodford, Michael (2002). "Imperfect Common Knowledge and the Effects of Monetary Policy," In Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps, ed. Philippe Aghion et al. Princeton and Oxford: Princeton University Press.

[32] Woodford, Michael (2003). Interest and Prices. Foundations of a Theory of Monetary Policy. Princeton and Oxford: Princeton University Press.

[33] Woodford, Michael (2009). "Information-Constrained State-Dependent Pricing," *Journal of Monetary Economics*, 56(S), 100-124.

Figure 1: Optimal Monetary Policy, Exogenous Information Structure

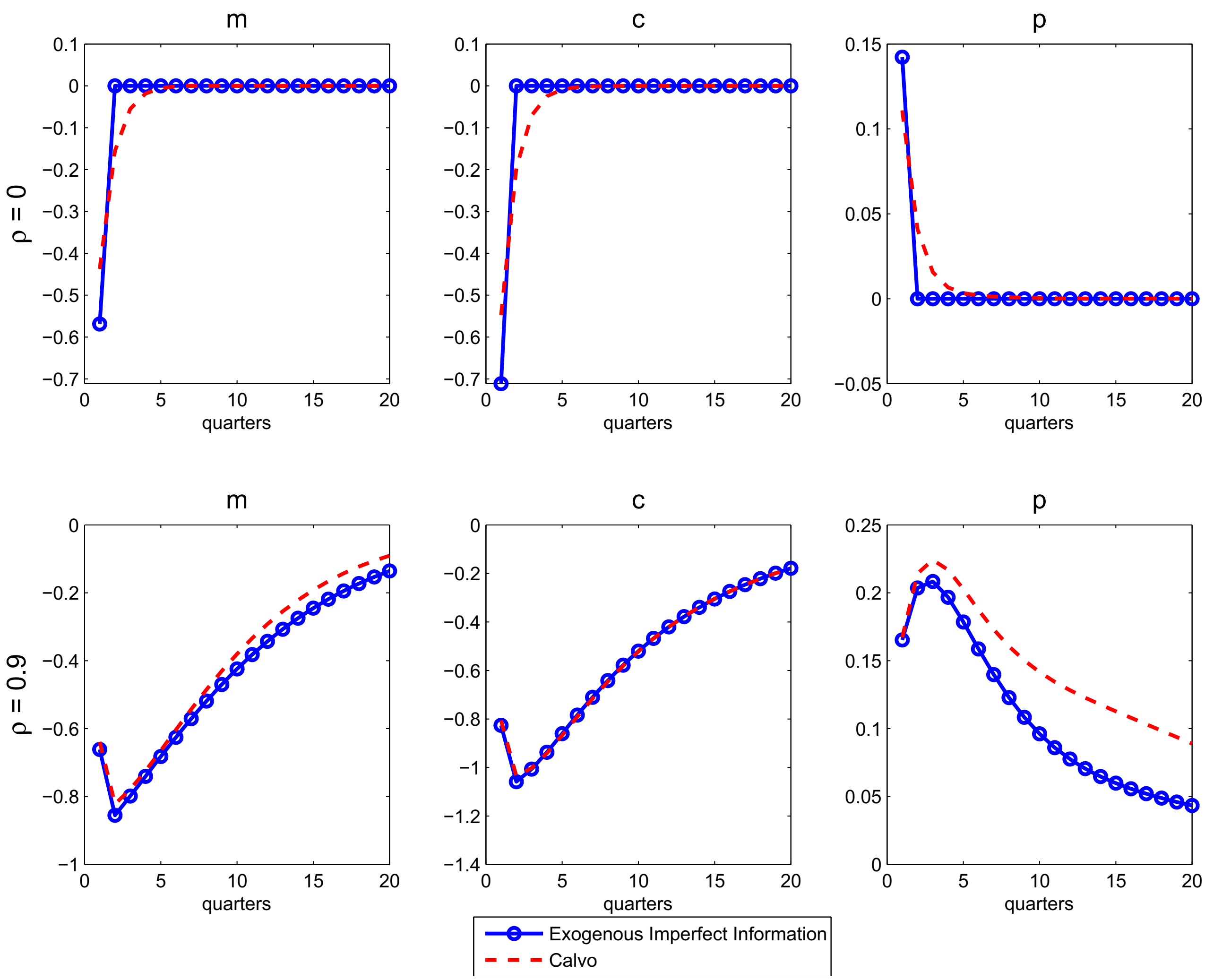

Figure 2: Equilibrium Attention as a Function of b

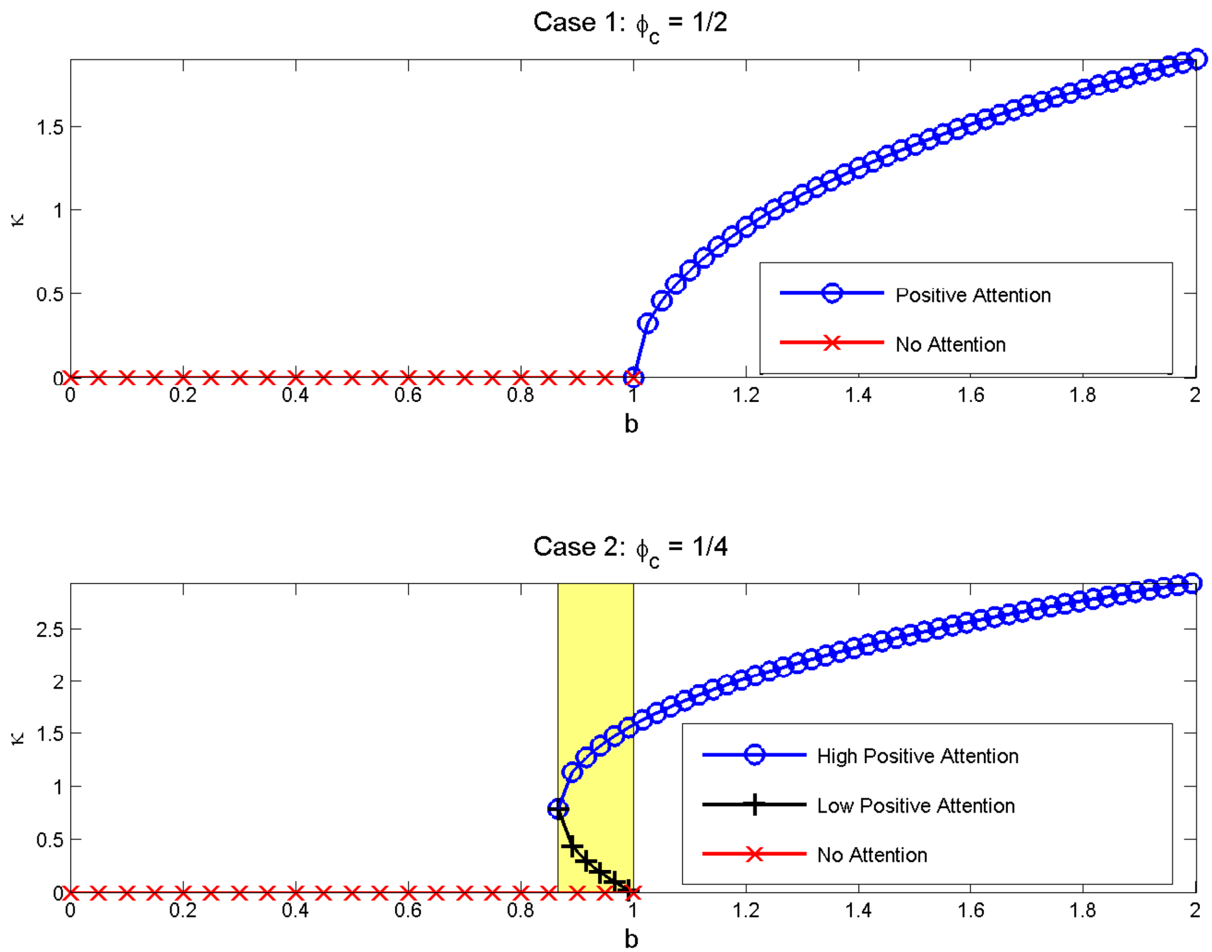

Figure 3: Optimal Monetary Policy, Endogenous Information Structure, iid Desired Markup

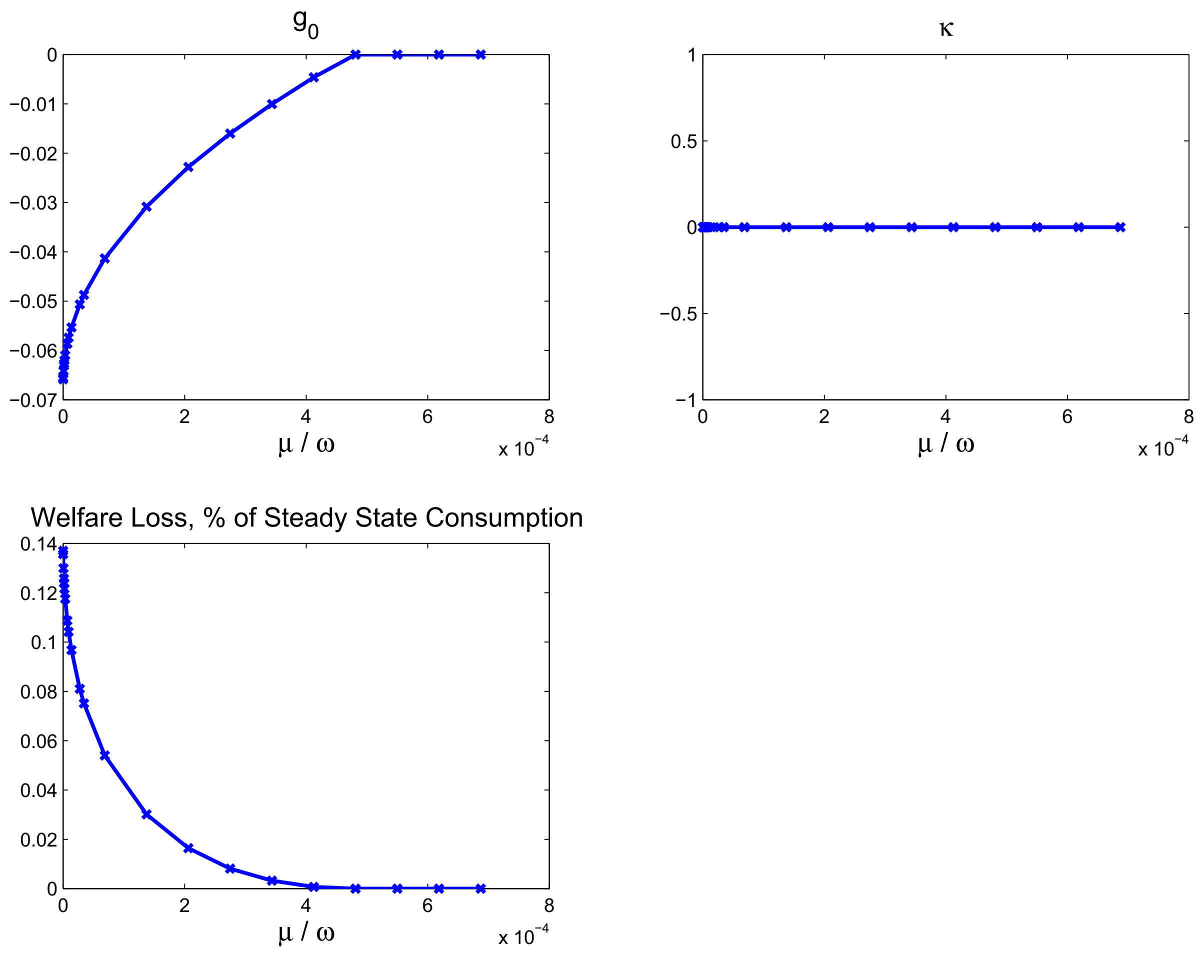

Figure 4: Optimal Monetary Policy, Endogenous Information Structure, Persistent Desired Markup

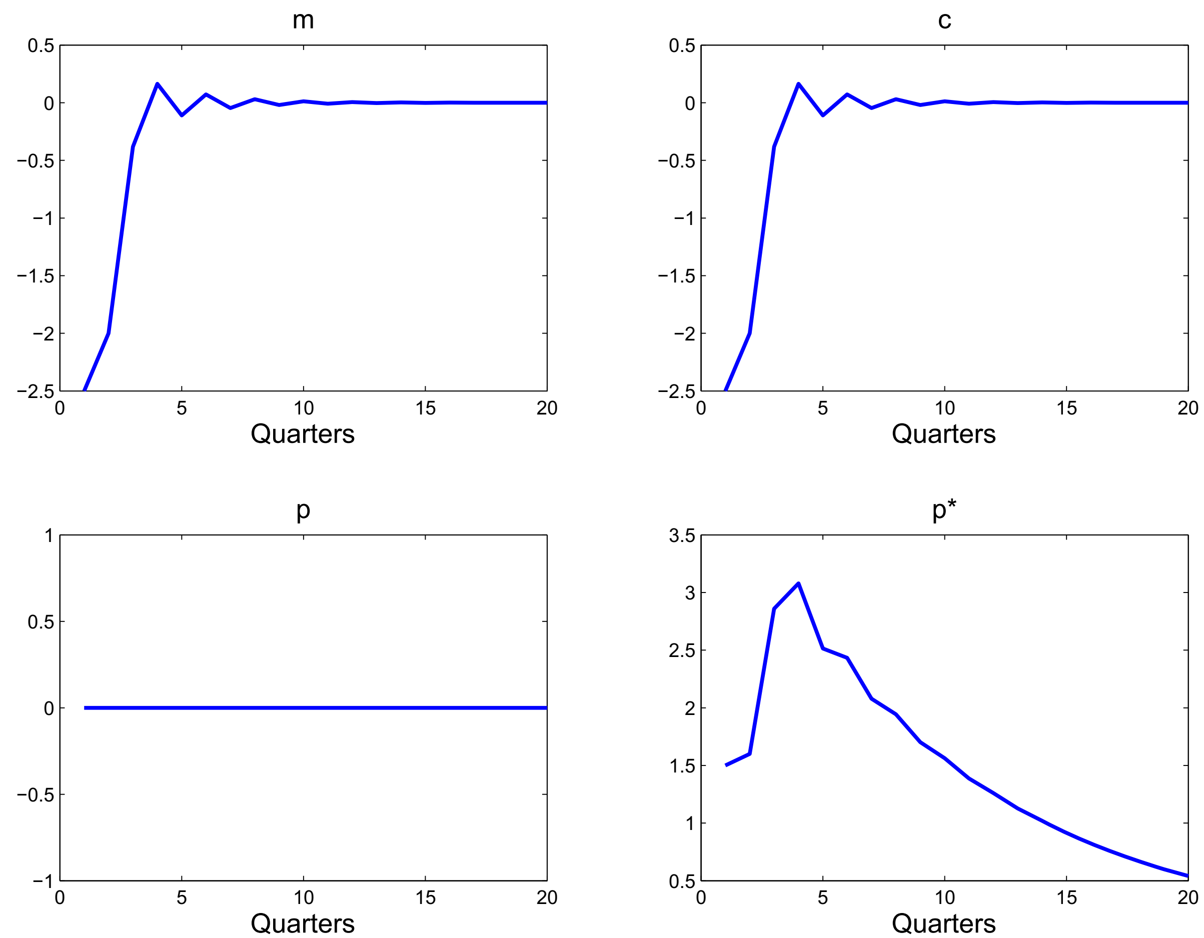

Figure 5: Welfare Losses under Endogenous Information for Alternative Policies