EFCE
7/9/08
11:30 am

Business Cycle Dynamics under Rational Inattention*

Bartosz Maćkowiak
European Central Bank and CEPR

Mirko Wiederholt
Northwestern University

First draft: June 2007. This draft: June 2008

Preliminary Version

Abstract

This paper develops a dynamic stochastic general equilibrium model with rational inattention. Decisionmakers have limited attention and choose the optimal allocation of their attention. We study the implications of rational inattention for business cycle dynamics. For example, we study how rational inattention affects the impulse responses of prices and quantities to monetary policy, aggregate technology and micro-level shocks. The impulse responses under rational inattention have several properties of empirical impulse response functions, e.g., (i) prices respond slowly to monetary policy shocks, (ii) prices respond faster to aggregate technology shocks, and (iii) prices respond very fast to disaggregate shocks. In addition, profit losses due to deviations of the actual price from the profit-maximizing price are an order of magnitude smaller than in the Calvo model that generates the same real effects.

*We thank Paco Buera, James Costain, Christian Hellwig and Giorgio Primiceri as well as seminar and conference participants at Bank of Canada, Chicago Fed, Columbia, DePaul, European Central Bank, European University Institute, ESSIM 2008, MIT, Minneapolis Fed, 2008 North American Winter Meeting of the Econometric Society, Riksbank, Stony Brook and UCSD for helpful comments. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the European Central Bank.

E-mail: bartosz.mackowiak@ecb.int and m-wiederholt@northwestern.edu.

1 Introduction

This paper develops a dynamic stochastic general equilibrium model with rational inattention. We model the idea that agents cannot attend perfectly to all available information. Therefore, the mapping between economic conditions and the price and quantity decisions taken by agents is not perfect. Decisionmakers make mistakes. However, decisionmakers try to minimize these mistakes.

The economy consists of households, firms and a government. Households supply differentiated types of labor, consume a variety of goods, and hold nominal government bonds. Households take wage setting and consumption decisions. Firms supply differentiated goods that are produced with the different types of labor. Firms take price setting and input mix decisions. The central bank sets the nominal interest rate according to a Taylor rule. In the model prices and wages are physically fully flexible. Limited attention is the only source of slow adjustment. We compute the impulse responses of all variables to monetary policy shocks, aggregate technology shocks and micro-level shocks under rational inattention.

Rational inattention means that decisionmakers have limited attention and choose the allocation of their attention. Following Sims (2003), we model decisionmakers' limited attention as a constraint on information flow, and we let decisionmakers choose how to satisfy this constraint. In other words, decisionmakers decide what to focus on. For example, decisionmakers decide how to allocate their attention across their different decision problems. Furthermore, decisionmakers decide how to attend to the different factors that may affect an optimal decision.

We find that rational inattention by decisionmakers in firms has the following implications. First, for our parameter values, rational inattention by decisionmakers in firms implies that the impulse response of the price level to monetary policy shocks resembles the impulse response in a Calvo model with an average price duration of 7.5 months. At the same time, prices respond fairly quickly to aggregate technology shocks and almost perfectly to micro-level shocks. The reason is the optimal allocation of attention. Decisionmakers in firms decide to pay little attention to monetary policy, about twice as much attention to aggregate technology, and a lot of attention to firm-specific conditions. Therefore, prices respond slowly to monetary policy shocks, but fairly quickly to aggregate technology shocks, and almost perfectly to micro-level shocks. Second, losses in profits due to deviations of the actual price from the profit-maximizing price are an order of magnitude smaller than in the Calvo model that generates the same real effects. In particular, losses in profits due to suboptimal price responses to aggregate conditions are 12 times smaller than in the Calvo model; and losses in profits due to suboptimal price responses to firm-specific conditions are 25 times smaller than in the Calvo model. One reason is the optimal allocation of attention: under rational inattention prices respond slowly to monetary policy shocks, but fairly quickly to aggregate technology shocks, and almost perfectly to micro-level shocks. In contrast, in the Calvo model prices respond slowly to all these shocks. The other reason is that under rational inattention deviations of the actual price from the profit-maximizing price are less likely to be extreme than in the Calvo model. This also reduces profit losses. Third, rational inattention on the side of decisionmakers in firms implies that firms produce with a factor mix that tends to deviate from the optimal factor mix.

This paper is related to two strands of literature: the literature on business cycle models with imperfect information (e.g., Lucas (1972), Woodford (2002), Mankiw and Reis (2002), and Lorenzoni (2008)) and the literature on rational inattention (e.g., Sims (2003, 2006), Luo (2008), Maćkowiak and Wiederholt (2008), Van Nieuwerburgh and Veldkamp (2008), and Woodford (2008)). The main difference to the existing literature on business cycle models with imperfect information is that in the model presented below agents choose the information structure. The information structure is the outcome of an optimization problem. The main difference to the existing literature on rational inattention is that we solve a dynamic stochastic general equilibrium model.

The paper is organized as follows. Section 2 describes all features of the economy apart from the decisionmakers' problem of allocating their attention. Section 3 describes the steady state of the non-stochastic version of the economy. In Section 4 we derive the objective of decisionmakers in firms when they choose how to allocate their attention. In Section 5 we derive the objective of households when they choose how to allocate their attention. Section 6 describes issues related to aggregation. Section 7 characterizes the solution of the model under perfect information. Section 8 shows numerical solutions of the model under rational inattention by decisionmakers in firms. Section 9 presents some results concerning the implications of adding rational inattention by households. Section 10 concludes.

2 Model

In this section we describe all features of the economy apart from the information structure. Afterwards we solve the model for different assumptions about the information structure: (i) perfect information and (ii) rational inattention.

2.1 Households

There are J households. Households supply differentiated types of labor, consume a variety of goods, and hold nominal government bonds.

Each household seeks to maximize the expected sum of discounted period utility. The discount factor is $\beta \in (0,1)$. The period utility function is given by

$$U\left(C_{jt}, L_{j1t}, \ldots, L_{jNt}\right) = \frac{C_{jt}^{1-\gamma} - 1}{1-\gamma} - \varphi \sum_{n=1}^{N} e^{\chi_{jnt}} \frac{L_{jnt}^{1+\psi}}{1+\psi}, \tag{1}$$

with[1]

$$C_{jt} = \left(\sum_{i=1}^{I} C_{ijt}^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}. \tag{2}$$

Here C_{jt} is composite consumption by household j in period t and C_{ijt} is consumption of good i by household j in period t. The household can consume I different goods. The parameter $\theta > 1$ is the elasticity of substitution between different goods. The parameter $\gamma > 0$ is the inverse of the intertemporal elasticity of substitution. We assume that households supply differentiated types of labor. Furthermore, we assume that each household supplies multiple types of labor. This assumption allows us to introduce labor-specific preference shocks that can be insured within the household. L_{jnt} denotes the supply of household j's nth type of labor in period t, and χ_{jnt} denotes a preference shock affecting the disutility of supplying this type of labor. We introduce labor-specific preference shocks to generate

[1] The assumption of a constant elasticity of substitution between varieties is only for ease of exposition. We have also done all the derivations using a general constant returns-to-scale aggregator.

variation in relative wage rates.[2] We let each household supply N types of labor to allow for risk sharing within the household. We will assume that N is large. The parameter $\varphi > 0$ affects the disutility of supplying labor. The parameter $\psi \geq 0$ is the inverse of the Frisch elasticity of labor supply.

Households can hold nominal government bonds. The flow budget constraint of household j in period t reads

$$\sum_{i=1}^{I} P_{it}C_{ijt} + B_{jt} = R_{t-1}B_{jt-1} + (1+\tau_w)\sum_{n=1}^{N} W_{jnt}L_{jnt} + \frac{D_t}{J} - \frac{T_t}{J}. \tag{3}$$

Here P_{it} is the price of good i in period t, B_{jt} are bond holdings by household j between periods t and $t+1$, R_{t-1} is the nominal interest rate on bond holdings between periods $t-1$ and t, τ_w is a wage subsidy, W_{jnt} is the nominal wage rate for household j's nth type of labor in period t, (D_t/J) is a pro-rata share of nominal aggregate profits, and (T_t/J) is a pro-rata share of nominal lump-sum taxes. We assume that all J households have the same initial bond holdings, $B_{j,-1} > 0$. Furthermore, we assume that bond holdings cannot be negative or zero, $B_{jt} > 0$. One can think of households holding an account. The account holds only nominal government bonds and the balance on the account has to be positive.

Every period each household chooses a consumption vector, $(C_{1jt}, \ldots, C_{Ijt})$, and a vector of nominal wage rates, $(W_{j1t}, \ldots, W_{jNt})$. Each household commits to supply any quantity of labor at the chosen nominal wage rates. Bond holdings, B_{jt}, then follow as a residual from the flow budget constraint (3) and the labor demand function derived below.

Households take as given all aggregate variables, all wage rates set by other households and all prices set by firms.

2.2 Firms

There are I firms in the economy. Firms supply differentiated goods that are produced with the different types of labor.

The technology of firm i is given by

$$Y_{it} = e^{a_t} e^{a_{it}} L_{it}^{\alpha}, \tag{4}$$

[2] Alternatively, one can introduce labor-specific productivity shocks to generate variation in relative wage rates. When we did that we obtained very similar results.

with

$$L_{it} = \left(\sum_{j=1}^{J} \sum_{n=1}^{N} L_{ijnt}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}. \tag{5}$$

Here Y_{it} is output of firm i in period t. L_{it} is composite labor input at firm i in period t and L_{ijnt} is input of household j's nth type of labor at firm i in period t. There are JN types of labor because there are J households each supplying N differentiated types of labor. The parameter $\eta > 1$ is the elasticity of substitution between different types of labor. The parameter $\alpha \in (0, 1]$ is the elasticity of output with respect to composite labor. Total factor productivity, $(e^{a_t} e^{a_{it}})$, has an aggregate component, e^{a_t}, and a firm-specific component, $e^{a_{it}}$.[3]

Nominal profits of firm i in period t equal

$$(1 + \tau_p) P_{it} Y_{it} - \sum_{j=1}^{J} \sum_{n=1}^{N} W_{jnt} L_{ijnt}, \tag{6}$$

where τ_p is a production subsidy.

Every period each firm chooses a price, P_{it}, and a factor mix, $\left(\hat{L}_{i11t}, \ldots, \hat{L}_{iJ(N-1)t}\right)$ with $\hat{L}_{ijnt} = (L_{ijnt}/L_{it})$. Each firm commits to supply any quantity of the good at the chosen price. The composite labor input, L_{it}, then follows from the production function (4) and the demand function derived below.

Firms take as given all aggregate variables, all prices set by other firms and all wage rates set by households.

2.3 Government

There is a monetary authority and a fiscal authority. Let $\Pi_t = (P_t/P_{t-1})$ denote inflation of a price index, P_t, that will be defined later, and let $Y_t = \left(\sum_{i=1}^{I} Y_{it}\right)$ denote aggregate output. The central bank sets the nominal interest rate according to the rule

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\rho_R} \left[\left(\frac{\Pi_t}{\Pi}\right)^{\phi_\pi} \left(\frac{Y_t}{Y}\right)^{\phi_y} \right]^{1-\rho_R} e^{\varepsilon_t^R}, \tag{7}$$

[3] The assumption that all types of labor appear in the labor aggregator (5) is only for ease of exposition. One can also use a labor aggregator in which a firm-specific subset of types of labor appears.

where ε_t^R is a monetary policy shock. Here R, Π and Y are the values of the nominal interest rate, inflation and aggregate output in the non-stochastic steady state. The policy parameters satisfy $\rho_R \in [0,1)$, $\phi_\pi > 1$ and $\phi_y \geq 0$.

The government budget constraint in period t reads

$$T_t + (B_t - B_{t-1}) = (R_{t-1} - 1)\, B_{t-1} + \tau_w \left(\sum_{j=1}^{J} \sum_{n=1}^{N} W_{jnt} L_{jnt} \right) + \tau_p \left(\sum_{i=1}^{I} P_{it} Y_{it} \right). \tag{8}$$

The government has to finance interest on nominal government bonds, the wage subsidy and the production subsidy. The government can collect lump-sum taxes or issue new government bonds.

We assume that the government sets the production subsidy, τ_p, and the wage subsidy, τ_w, so as to correct the distortions that arise from firms' market power in the goods market and households' market power in the labor market. In particular, we assume that

$$1 + \tau_p = \frac{\vartheta}{\vartheta - 1}, \tag{9}$$

and

$$1 + \tau_w = \frac{\zeta}{\zeta - 1}, \tag{10}$$

where ϑ denotes the price elasticity of demand and ζ denotes the wage elasticity of labor demand.[4]

2.4 Shocks

There are four types of shocks in the economy: aggregate technology shocks, monetary policy shocks, firm-specific productivity shocks and labor-specific preference shocks. We assume that, for all i and jn, the processes $\{a_t\}$, $\{\varepsilon_t^R\}$, $\{a_{it}\}$ and $\{\chi_{jnt}\}$ are independent. Furthermore, we assume that firm-specific productivity processes, $\{a_{it}\}$, are independent across firms and labor-specific preference shocks, $\{\chi_{jnt}\}$, are independent across types of labor. In addition, we assume that I and N are sufficiently large so that

$$\frac{1}{I} \sum_{i=1}^{I} a_{it} = 0, \tag{11}$$

[4] When households have perfect information $\vartheta = \theta$, but in general ϑ may differ from θ. Similarly, when firms have perfect information $\zeta = \eta$, but in general ζ may differ from η.

and

$$\frac{1}{N}\sum_{n=1}^{N}\chi_{jnt}=0. \tag{12}$$

Finally, we assume that all exogenous processes are stationary Gaussian processes with mean zero. In the following, we denote the period t innovation to a_t, a_{it} and χ_{jnt} by ε_t^A, ε_{it}^I and ε_{jnt}^{χ}, respectively.

3 Non-stochastic steady state

Before studying the economy outlined in Section 2 that is hit by shocks, we characterize the non-stochastic steady state. The non-stochastic steady state is defined as the solution of the non-stochastic version of the economy that has the property that all quantities, all relative prices, the nominal interest rate and inflation are constant over time. In the following, variables without the subscript t denote values in the non-stochastic steady state.

In this section, let P_t denote the following price index

$$P_t=\left(\sum_{i=1}^{I}P_{it}^{1-\theta}\right)^{\frac{1}{1-\theta}}, \tag{13}$$

and let W_t denote the following wage index

$$W_t=\left(\sum_{j=1}^{J}\sum_{n=1}^{N}W_{jnt}^{1-\eta}\right)^{\frac{1}{1-\eta}}. \tag{14}$$

Throughout the paper, let $\hat{P}_{it}$ denote the relative price of good i, let $\hat{W}_{jnt}$ denote the relative wage rate for type jn labor, let $\tilde{W}_{jnt}$ denote the real wage rate for type jn labor, and let $\tilde{W}_t$ denote the real wage index. Formally,

$$\hat{P}_{it} = \frac{P_{it}}{P_t}, \tag{15}$$

$$\hat{W}_{jnt} = \frac{W_{jnt}}{W_t}, \tag{16}$$

$$\tilde{W}_{jnt} = \frac{W_{jnt}}{P_t}, \tag{17}$$

$$\tilde{W}_t = \frac{W_t}{P_t}. \tag{18}$$

It is straightforward to show that in the non-stochastic steady state the households' first-order conditions read

$$\frac{R}{\Pi} = \frac{1}{\beta}, \tag{19}$$

$$\frac{C_{ij}}{C_j} = \hat{P}_i^{-\theta}, \tag{20}$$

and

$$\tilde{W}_{jn} = \frac{1}{1+\tau_w}\frac{\eta}{\eta-1}\varphi\left(\hat{W}_{jn}^{-\eta}L\right)^{\psi}C_j^{\gamma}, \tag{21}$$

with

$$L = \sum_{i=1}^{I} L_i. \tag{22}$$

Furthermore, the firms' first-order conditions read

$$\frac{L_{ijn}}{L_i} = \hat{W}_{jn}^{-\eta}, \tag{23}$$

and

$$\hat{P}_i = \frac{1}{1+\tau_p}\frac{\theta}{\theta-1}\tilde{W}\frac{1}{\alpha}\left(\hat{P}_i^{-\theta}C\right)^{\frac{1}{\alpha}-1}, \tag{24}$$

with

$$C = \sum_{j=1}^{J} C_j. \tag{25}$$

Since all households face the same decision problem in the non-stochastic version of the model, all households choose the same level of composite consumption. It follows from the definition of aggregate composite consumption (25) that

$$C = JC_j. \tag{26}$$

The households' wage setting equation (21) then implies that the wage rate is the same for all different types of labor. Therefore, each firm hires the different types of labor in equal amounts. It follows from the definition of the wage index (14) and the labor aggregator (5) that

$$\hat{W}_{jn}^{1-\eta} = \left(\frac{L_{ijn}}{L_i}\right)^{\frac{\eta-1}{\eta}} = \frac{1}{JN}. \tag{27}$$

Similarly, the firms' price setting equation (24) implies that all firms set the same price. Therefore, each household consumes the different goods in equal amounts. It follows from the definition of the price index (13), the consumption aggregator (2), the definition of aggregate output and the definition of aggregate labor input (22) that

$$\hat{P}_i^{1-\theta} = \left(\frac{C_{ij}}{C_j}\right)^{\frac{\theta-1}{\theta}} = \frac{Y_i}{Y} = \frac{L_i}{L} = \frac{1}{I}. \tag{28}$$

Finally, note that neither the nominal interest rate, R, nor the inflation rate, Π, is uniquely determined in the non-stochastic steady state. Only the real interest rate, (R/Π), is uniquely determined in the non-stochastic steady state. Fortunately, in the following the values of R and Π will not matter. Only the value of (R/Π) will matter.

4 Derivation of the firms' objective

First, we guess that the demand function for good i has the form

$$C_{it} = \varsigma \left(\frac{P_{it}}{P_t}\right)^{-\vartheta} C_t, \tag{29}$$

where $\varsigma > 0$ and $\vartheta > 1$ are undetermined coefficients satisfying $\varsigma \hat{P}_i^{-\vartheta} C = \hat{P}_i^{-\theta} C$,

$$C_t = \sum_{j=1}^{J} C_{jt}, \tag{30}$$

and P_t is some price index satisfying

$$1 = \sum_{i=1}^{I} d\left(\frac{P_{it}}{P_t}\right), \tag{31}$$

where d is a twice continuously differentiable function. We will always make assumptions to ensure that the guess (29)-(31) is correct. When households have perfect information, $\varsigma = 1$, $\vartheta = \theta$ and the price index P_t is given by equation (13). When households have limited attention, we will need to add assumptions concerning the exogenous processes to ensure that the demand function still has the form (29)-(31).

Second, given a demand function, we can derive the profit function. The profit function is obtained by substituting the production function (4), the labor aggregator (5) and the demand function (29) into the expression for nominal profits (6). To do this it will be helpful to rewrite the expression for nominal profits (6) as

$$(1+\tau_p) P_{it} Y_{it} - L_{it} \left(\sum_{j=1}^{J} \sum_{n=1}^{N} W_{jnt} \hat{L}_{ijnt} \right), \tag{32}$$

where $\hat{L}_{ijnt} = (L_{ijnt}/L_{it})$. Here we have simply expressed the wage bill as the product of the composite labor input and the wage bill per unit of composite labor input. Rearranging the production function (4) yields

$$L_{it} = \left(\frac{Y_{it}}{e^{a_t} e^{a_{it}}}\right)^{\frac{1}{\alpha}}. \tag{33}$$

Rearranging the labor aggregator (5) yields

$$1 = \sum_{j=1}^{J} \sum_{n=1}^{N} \hat{L}_{ijnt}^{\frac{\eta-1}{\eta}},$$

or equivalently

$$\hat{L}_{iJNt} = \left(1 - \sum_{jn \neq JN} \hat{L}_{ijnt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}. \tag{34}$$

Substituting equations (33)-(34) and the demand function (29) into the expression for nominal profits (32) yields the profit function

$$\begin{aligned}
&(1+\tau_p) P_{it} \varsigma \left(\frac{P_{it}}{P_t}\right)^{-\vartheta} C_t \\
&- \left[\frac{\varsigma \left(\frac{P_{it}}{P_t}\right)^{-\vartheta} C_t}{e^{a_t} e^{a_{it}}}\right]^{\frac{1}{\alpha}} \left[\sum_{jn \neq JN} W_{jnt} \hat{L}_{ijnt} + W_{JNt} \left(1 - \sum_{jn \neq JN} \hat{L}_{ijnt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\right].
\end{aligned} \tag{35}$$

Nominal profits of firm i in period t depend on variables that the firm chooses, P_{it} and $\hat{L}_{i11t}, \ldots, \hat{L}_{iJ(N-1)t}$, as well as variables that the firm takes as given, P_t, C_t, e^{a_t}, $e^{a_{it}}$ and $W_{11t}, \ldots, W_{JNt}$.

Third, we assume that the firm chooses the allocation of attention in period $t = -1$ so as to maximize the expected sum of discounted profits. We assume that profits in period t are discounted with the following stochastic discount factor:

$$Q_{-1,t} = \beta^t \lambda (C_{1t}, \ldots, C_{Jt}) \frac{1}{P_t}, \tag{36}$$

where λ is a twice continuously differentiable function that has the property

$$\lambda (C_1, \ldots, C_J) = C_j^{-\gamma}, \tag{37}$$

and the price index P_t is the price index that appears in the demand function (29). For example, if the function λ is a weighted average of the marginal utility of the different households (i.e., $\lambda\left(C_{1t},\ldots,C_{Jt}\right)=\lambda_1 C_{1t}^{-\gamma}+\ldots+\lambda_J C_{Jt}^{-\gamma}$ with $\lambda_j\geq 0$ and $\sum_{j=1}^{J}\lambda_j=1$), then equation (37) is satisfied because all households have the same marginal utility in the non-stochastic steady state. The objective of firm i in period $t=-1$ then reads

$$E_{i,-1}\sum_{t=0}^{\infty}\beta^t\lambda\left(C_{1t},\ldots,C_{Jt}\right)\left[\begin{array}{c}\left(1+\tau_p\right)\varsigma\hat{P}_{it}^{1-\vartheta}C_t\\ -\left(\frac{\varsigma\hat{P}_{it}^{-\vartheta}C_t}{e^{a_t}e^{a_{it}}}\right)^{\frac{1}{\alpha}}\cdot\\ \left[\sum_{jn\neq JN}\tilde{W}_{jnt}\hat{L}_{ijnt}+\tilde{W}_{JNt}\left(1-\sum_{jn\neq JN}\hat{L}_{ijnt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\right]\end{array}\right],\quad(38)$$

where $\hat{P}_{it}=\left(P_{it}/P_t\right)$ is the relative price of good i, $\tilde{W}_{jnt}=\left(W_{jnt}/P_t\right)$ is the real wage rate for type jn labor, and $E_{i,-1}$ is the expectation operator conditioned on information of firm i in period $t=-1$.

Fourth, one can express objective (38) in terms of log-deviations from the non-stochastic steady state. In the following, let small variables denote log-deviations from the non-stochastic steady state, e.g., $c_t=\ln\left(C_t/C\right)$. Using the fact that $C_t=Ce^{c_t}$ yields the following expression for the objective of firm i in period $t=-1$

$$E_{i,-1}\sum_{t=0}^{\infty}\beta^t\lambda\left(C_1e^{c_{1t}},\ldots,C_Je^{c_{Jt}}\right)\left[\begin{array}{c}\left(1+\tau_p\right)\varsigma\hat{P}_{i}^{1-\vartheta}Ce^{(1-\vartheta)\hat{p}_{it}+c_t}\\ -\left(\frac{\varsigma\hat{P}_{i}^{-\vartheta}Ce^{-\vartheta\hat{p}_{it}+c_t}}{e^{a_t}e^{a_{it}}}\right)^{\frac{1}{\alpha}}\cdot\\ \frac{\tilde{W}}{JN}\left[\sum_{jn\neq JN}e^{\tilde{w}_{jnt}+\hat{l}_{ijnt}}+e^{\tilde{w}_{JNt}}\left(JN-\sum_{jn\neq JN}e^{\frac{\eta-1}{\eta}\hat{l}_{ijnt}}\right)^{\frac{\eta}{\eta-1}}\right]\end{array}\right],\quad(39)$$

where $\tilde{W}$ is the value of the wage index (14) in the non-stochastic steady state. Here we have used equations (23) and (27).

Proposition 1 *Let f denote the functional inside the expectation operator in expression (39). Let $\tilde{f}$ denote the second-order Taylor approximation of f at the non-stochastic steady state. Let x_t and z_t denote the following vectors*

$$x_t=\left(\begin{array}{cccc}\hat{p}_{it} & \hat{l}_{i11t} & \cdots & \hat{l}_{iJ(N-1)t}\end{array}\right)',\quad(40)$$

$$z_t=\left(\begin{array}{cccccccc}c_{1t} & \cdots & c_{Jt} & c_t & a_t & a_{it} & \tilde{w}_{11t} & \cdots & \tilde{w}_{JNt}\end{array}\right)'.\quad(41)$$

Suppose that

$$\forall t : E_{i,-1}\left|\begin{pmatrix} x_t \\ z_t \\ 1 \end{pmatrix}\begin{pmatrix} x_t \\ z_t \\ 1 \end{pmatrix}'\right| < \infty, \tag{42}$$

and

$$\lim_{t\to\infty}\beta^t E_{i,-1}\left|\begin{pmatrix} x_t \\ z_t \\ 1 \end{pmatrix}\begin{pmatrix} x_t \\ z_t \\ 1 \end{pmatrix}'\right| = 0. \tag{43}$$

Then

$$\begin{aligned} & E_{i,-1}\left[\tilde{f}\left(x_0, z_0, x_1, z_1, \ldots\right)\right] - E_{i,-1}\left[\tilde{f}\left(x_0^*, z_0, x_1^*, z_1, \ldots\right)\right] \\ = \ & \sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}\left(x_t - x_t^*\right)' H_x \left(x_t - x_t^*\right)\right], \end{aligned} \tag{44}$$

where

$$H_x = -C_j^{-\gamma}\tilde{W}L_i\begin{bmatrix} \frac{\vartheta}{\alpha}\left(1+\frac{1-\alpha}{\alpha}\vartheta\right) & 0 & \cdots & \cdots & 0 \\ 0 & \frac{2}{\eta JN} & \frac{1}{\eta JN} & \cdots & \frac{1}{\eta JN} \\ \vdots & \frac{1}{\eta JN} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1}{\eta JN} \\ 0 & \frac{1}{\eta JN} & \cdots & \frac{1}{\eta JN} & \frac{2}{\eta JN} \end{bmatrix}, \tag{45}$$

and x_t^* is given by

$$\hat{p}_{it}^* = \frac{\frac{1-\alpha}{\alpha}}{1+\frac{1-\alpha}{\alpha}\vartheta}c_t - \frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\vartheta}\left(a_t + a_{it}\right) + \frac{1}{1+\frac{1-\alpha}{\alpha}\vartheta}\left(\frac{1}{JN}\sum_{j=1}^{J}\sum_{n=1}^{N}\tilde{w}_{jnt}\right), \tag{46}$$

and

$$\hat{l}_{ijnt}^* = -\eta\left(\tilde{w}_{jnt} - \frac{1}{JN}\sum_{j=1}^{J}\sum_{n=1}^{N}\tilde{w}_{jnt}\right). \tag{47}$$

Proof. See Appendix A. ■

5 Derivation of the households' objective

First, we guess that the demand function for type jn labor has the form

$$L_{jnt} = \xi\left(\frac{W_{jnt}}{W_t}\right)^{-\zeta} L_t, \tag{48}$$

where $\xi > 0$ and $\zeta > 1$ are undetermined coefficients satisfying $\xi \hat{W}_{jn}^{-\zeta} L = \hat{W}_{jn}^{-\eta} L$,

$$L_t = \sum_{i=1}^{I} L_{it}, \tag{49}$$

and W_t is some wage index satisfying

$$1 = \sum_{j=1}^{J} \sum_{n=1}^{N} h\left(\frac{W_{jnt}}{W_t}\right), \tag{50}$$

where h is a twice continuously differentiable function. We will always make assumptions to ensure that the guess (48)-(50) is correct. When firms have perfect information, $\xi = 1$, $\zeta = \eta$ and the wage index W_t is given by equation (14). When firms have limited attention, we will need to add assumptions concerning the exogenous processes to ensure that the labor demand function still has the form (48)-(50).

Second, just like we derived the profit function for firms, we will now derive an objective for households that incorporates all the constraints. One can write the flow budget constraint (3) as

$$C_{jt}\left(\sum_{i=1}^{I} P_{it}\hat{C}_{ijt}\right) + B_{jt} = R_{t-1}B_{jt-1} + (1+\tau_w)\sum_{n=1}^{N} W_{jnt}L_{jnt} + \frac{D_t}{J} - \frac{T_t}{J},$$

where $\hat{C}_{ijt} = (C_{ijt}/C_{jt})$. Here we have simply expressed consumption expenditure as the product of composite consumption and expenditure per unit of composite consumption. Rearranging the last equation yields

$$C_{jt} = \frac{R_{t-1}B_{jt-1} - B_{jt} + (1+\tau_w)\sum_{n=1}^{N} W_{jnt}L_{jnt} + \frac{D_t}{J} - \frac{T_t}{J}}{\sum_{i=1}^{I} P_{it}\hat{C}_{ijt}}.$$

Dividing the numerator and the denominator on the right-hand side by P_t, where P_t is some price index, yields

$$C_{jt} = \frac{\frac{R_{t-1}}{\Pi_t}\tilde{B}_{jt-1} - \tilde{B}_{jt} + (1+\tau_w)\sum_{n=1}^{N} \tilde{W}_{jnt}L_{jnt} + \frac{\tilde{D}_t}{J} - \frac{\tilde{T}_t}{J}}{\sum_{i=1}^{I} \hat{P}_{it}\hat{C}_{ijt}}, \tag{51}$$

where $\hat{P}_{it} = (P_{it}/P_t)$ is the relative price of good i, $\tilde{W}_{jnt} = (W_{jnt}/P_t)$ is the real wage rate for type jn labor, $\tilde{B}_{jt} = (B_{jt}/P_t)$ are real bond holdings by household j, $\tilde{D}_t = (D_t/P_t)$ are real aggregate profits, $\tilde{T}_t = (T_t/P_t)$ are real lump-sum taxes, and $\Pi_t = (P_t/P_{t-1})$ is inflation. Rearranging the consumption aggregator (2) yields

$$1 = \sum_{i=1}^{I} \hat{C}_{ijt}^{\frac{\theta-1}{\theta}},$$

or equivalently

$$\hat{C}_{Ijt} = \left(1 - \sum_{i=1}^{I-1} \hat{C}_{ijt}^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}. \tag{52}$$

Substituting equations (51)-(52) and the labor demand function (48) into the period utility function (1) yields the following expression for period utility of household j in period t

$$\frac{1}{1-\gamma}\left(\frac{\frac{R_{t-1}}{\Pi_t}\tilde{B}_{jt-1} - \tilde{B}_{jt} + (1+\tau_w)\sum_{n=1}^{N}\tilde{W}_{jnt}\xi\left(\frac{\tilde{W}_{jnt}}{\tilde{W}_t}\right)^{-\zeta}L_t + \frac{\tilde{D}_t}{J} - \frac{\tilde{T}_t}{J}}{\sum_{i=1}^{I-1}\hat{P}_{it}\hat{C}_{ijt} + \hat{P}_{It}\left(1 - \sum_{i=1}^{I-1}\hat{C}_{ijt}^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}}\right)^{1-\gamma} - \frac{1}{1-\gamma}$$
$$-\varphi\sum_{n=1}^{N} e^{\chi_{jnt}}\frac{1}{1+\psi}\left[\xi\left(\frac{\tilde{W}_{jnt}}{\tilde{W}_t}\right)^{-\zeta}L_t\right]^{1+\psi}, \tag{53}$$

where $\tilde{W}_t = (W_t/P_t)$ is the wage index divided by the price index. Here we have used the fact that $(W_{jnt}/W_t) = \left(\tilde{W}_{jnt}/\tilde{W}_t\right)$. The only constraint that we have not incorporated yet in the expression for period utility is the constraint that bond holdings must be positive, $\tilde{B}_{jt} > 0$. We will incorporate this constraint below by expressing all variables in terms of log-deviations from the non-stochastic steady state.

Third, expressing all variables in terms of log-deviations from the non-stochastic steady state yields the following expression for period utility of household j in period t

$$\frac{1}{1-\gamma}\left(\frac{\frac{Re^{r_{t-1}}}{\Pi e^{\pi_t}}\tilde{B}_j e^{\tilde{b}_{jt-1}}-\tilde{B}_j e^{\tilde{b}_{jt}}+(1+\tau_w)\sum_{n=1}^{N}\tilde{W}_{jn}e^{\tilde{w}_{jnt}}\xi\left(\frac{\tilde{W}_{jn}e^{\tilde{w}_{jnt}}}{\tilde{W}e^{\tilde{w}_t}}\right)^{-\zeta}Le^{l_t}+\frac{\tilde{D}e^{\tilde{d}_t}}{J}-\frac{\tilde{T}e^{\tilde{t}_t}}{J}}{\frac{1}{I}\left[\sum_{i=1}^{I-1}e^{\hat{p}_{it}+\hat{c}_{ijt}}+e^{\hat{p}_{It}}\left(I-\sum_{i=1}^{I-1}e^{\frac{\theta-1}{\theta}\hat{c}_{ijt}}\right)^{\frac{\theta}{\theta-1}}\right]}\right)^{1-\gamma}$$

$$-\frac{1}{1-\gamma}-\varphi\sum_{n=1}^{N}e^{\chi_{jnt}}\frac{1}{1+\psi}\left[\xi\left(\frac{\tilde{W}_{jn}e^{\tilde{w}_{jnt}}}{\tilde{W}e^{\tilde{w}_t}}\right)^{-\zeta}Le^{l_t}\right]^{1+\psi}. \tag{54}$$

Here we have used equations (20) and (28).

In the following, let $(\tilde{w}_{jnt})$, $(\hat{p}_{it})$ and (χ_{jnt}) denote the $1\times N$ vector of real wage rates, the $1\times I$ vector of relative goods prices and the $1\times N$ vector of preference shocks, respectively. Furthermore, let ω_B, ω_W, ω_D and ω_T denote the following ratios in the non-stochastic steady state

$$\begin{pmatrix}\omega_B & \omega_W & \omega_D & \omega_T\end{pmatrix}=\begin{pmatrix}\frac{\tilde{B}_j}{C_j} & \frac{(1+\tau_w)\sum_{n=1}^{N}\tilde{W}_{jn}L_{jn}}{C_j} & \frac{\frac{\tilde{D}}{J}}{C_j} & \frac{\frac{\tilde{T}}{J}}{C_j}\end{pmatrix}.$$

Proposition 2 *Let g denote the functional that is obtained by multiplying the expression for period utility (54) by β^t and summing over all t from zero to infinity. Let $\tilde{g}$ denote the second-order Taylor approximation of g at the non-stochastic steady state. Let x_t and z_t denote the following vectors*

$$x_t = \begin{pmatrix}\tilde{b}_{jt} & (\tilde{w}_{jnt}) & \hat{c}_{1jt} & \cdots & \hat{c}_{I-1jt}\end{pmatrix}', \tag{55}$$

$$z_t = \begin{pmatrix}r_{t-1} & \pi_t & \tilde{w}_t & l_t & \tilde{d}_t & \tilde{t}_t & (\hat{p}_{it}) & (\chi_{jnt})\end{pmatrix}'. \tag{56}$$

Suppose that for $\tau=0,1$

$$\forall t: E_{j,-1}\left|\begin{pmatrix}x_t\\ z_t\\ 1\end{pmatrix}\begin{pmatrix}x_{t+\tau}\\ z_{t+\tau}\\ 1\end{pmatrix}'\right|<\infty, \tag{57}$$

and

$$\lim_{t\to\infty}\beta^t E_{j,-1}\left|\begin{pmatrix}x_t\\ z_t\\ 1\end{pmatrix}\begin{pmatrix}x_{t+\tau}\\ z_{t+\tau}\\ 1\end{pmatrix}'\right|=0. \tag{58}$$

Then

$$\begin{aligned} & E_{j,-1}\left[\tilde{g}\left(x_0, z_0, x_1, z_1, \ldots\right)\right] - E_{j,-1}\left[\tilde{g}\left(x_0^*, z_0, x_1^*, z_1, \ldots\right)\right] \\ = & \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2}\left(x_t - x_t^*\right)' H_0 \left(x_t - x_t^*\right) + \left(x_t - x_t^*\right)' H_1 \left(x_{t+1} - x_{t+1}^*\right)\right], \end{aligned} \tag{59}$$

where

$$H_0 = -\gamma C_j^{1-\gamma} \begin{bmatrix} \omega_B^2 + \frac{\omega_B^2}{\beta} & \frac{\omega_B \omega_W (\zeta-1)}{N} & \cdots & \frac{\omega_B \omega_W (\zeta-1)}{N} & 0 & \cdots & 0 \\ \frac{\omega_B \omega_W (\zeta-1)}{N} & \frac{\omega_W(\zeta-1)}{N}\left(\frac{\omega_W(\zeta-1)}{N} + \frac{1+\psi\zeta}{\gamma}\right) & \cdots & \left(\frac{\omega_W(\zeta-1)}{N}\right)^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\omega_B \omega_W (\zeta-1)}{N} & \left(\frac{\omega_W(\zeta-1)}{N}\right)^2 & \cdots & \frac{\omega_W(\zeta-1)}{N}\left(\frac{\omega_W(\zeta-1)}{N} + \frac{1+\psi\zeta}{\gamma}\right) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \frac{2}{\gamma\theta I} & \cdots & \frac{1}{\gamma\theta I} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \frac{1}{\gamma\theta I} & \cdots & \frac{2}{\gamma\theta I} \end{bmatrix}, \tag{60}$$

$$H_1 = \gamma C_j^{1-\gamma} \begin{bmatrix} \omega_B^2 & \frac{\omega_B \omega_W (\zeta-1)}{N} & \cdots & \frac{\omega_B \omega_W (\zeta-1)}{N} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{61}$$

and x_t^* *is given by*

$$\begin{aligned} \tilde{b}_{jt}^* = & \frac{1}{\beta}\left(r_{t-1} - \pi_t + \tilde{b}_{jt-1}^*\right) + \frac{\omega_W}{\omega_B}\frac{1}{N}\sum_{n=1}^{N}\left[\tilde{w}_{jnt}^* - \zeta\left(\tilde{w}_{jnt}^* - \tilde{w}_t\right) + l_t\right] \\ & + \frac{\omega_D}{\omega_B}\tilde{d}_t - \frac{\omega_T}{\omega_B}\tilde{t}_t - \frac{1}{\omega_B}c_{jt}^*, \end{aligned} \tag{62}$$

$$c_{jt}^* = E_t\left[-\frac{1}{\gamma}\left(r_t - \pi_{t+1}\right) + c_{jt+1}^*\right], \tag{63}$$

$$\tilde{w}_{jnt}^* = \frac{\gamma}{1+\psi\zeta}c_{jt}^* + \frac{\psi}{1+\psi\zeta}\left(\zeta\tilde{w}_t + l_t\right) + \frac{1}{1+\psi\zeta}\chi_{jnt}, \tag{64}$$

and

$$\hat{c}_{ijt}^* = -\theta\hat{p}_{it}, \tag{65}$$

as well as the requirement that x_t^* *satisfies (57)-(58).*

Proof. See Appendix B. ■

6 Aggregation

In the following, we will work with log-linearized equations for the aggregate variables. Log-linearizing the equations for aggregate output, aggregate composite consumption (30) and aggregate composite labor input (49) yields

$$y_t = \frac{1}{I}\sum_{i=1}^{I} y_{it}, \tag{66}$$

$$c_t = \frac{1}{J}\sum_{j=1}^{J} c_{jt}, \tag{67}$$

and

$$l_t = \frac{1}{I}\sum_{i=1}^{I} l_{it}. \tag{68}$$

Log-linearizing the equation for the price index (31) yields

$$0 = \sum_{i=1}^{I} \hat{p}_{it},$$

or equivalently

$$p_t = \frac{1}{I}\sum_{i=1}^{I} p_{it}. \tag{69}$$

Log-linearizing the equation for the wage index (50) yields

$$0 = \sum_{j=1}^{J}\sum_{n=1}^{N} \hat{w}_{jnt},$$

or equivalently

$$w_t = \frac{1}{JN}\sum_{j=1}^{J}\sum_{n=1}^{N} w_{jnt}. \tag{70}$$

Finally, the production function (4) and the monetary policy rule (7) are already log-linear

$$y_{it} = a_t + a_{it} + \alpha l_{it}, \tag{71}$$

and

$$r_t = \rho_R r_{t-1} + (1-\rho_R)\left(\phi_\pi \pi_t + \phi_y y_t\right) + \varepsilon_t^R. \tag{72}$$

7 Case 1: Firms and households have perfect information

Proposition 3 *Suppose that firms and households know the realization of all variables up to and including period t. Then each firm chooses a price and a factor mix satisfying (46)-(47), and each household chooses a consumption vector and a vector of nominal wage rates satisfying (62)-(65). A solution to the system of equations $y_{it} = c_{it} = (1/J)\sum_{j=1}^{J} c_{ijt}$, (11)-(12), (46)-(47), (62)-(65) and (66)-(71) with a non-explosive bond sequence for each household satisfies:*

$$
\begin{aligned}
c_t &= \frac{1+\psi}{1-\alpha+\alpha\gamma+\psi} a_t, && (73)\\
l_t &= \frac{1-\gamma}{1-\alpha+\alpha\gamma+\psi} a_t, && (74)\\
\tilde{w}_t &= \frac{\gamma+\psi}{1-\alpha+\alpha\gamma+\psi} a_t, && (75)\\
r_t - E_t\left[\pi_{t+1}\right] &= \gamma\frac{1+\psi}{1-\alpha+\alpha\gamma+\psi} E_t\left[a_{t+1}-a_t\right], && (76)
\end{aligned}
$$

as well as

$$
\begin{aligned}
\hat{p}_{it} &= -\frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\theta} a_{it}, && (77)\\
\hat{c}_{ijt} &= \theta\frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\theta} a_{it}, && (78)\\
\tilde{w}_{jnt} - \tilde{w}_t &= \frac{1}{1+\psi\eta}\chi_{jnt}, && (79)\\
\hat{l}_{ijnt} &= -\eta\frac{1}{1+\psi\eta}\chi_{jnt}. && (80)
\end{aligned}
$$

Proof. *See Appendix C.* ■

Under perfect information aggregate output, aggregate consumption, the aggregate labor input, the real wage index and the real interest rate only depend on aggregate technology. Monetary policy has no real effects.

The nominal interest rate and inflation then follow from the monetary policy rule (72) and the real interest rate (76). Since $(1-\rho_R)\phi_\pi > 0$ and $(1-\rho_R)\phi_\pi + \rho_R > 1$, the equilibrium paths of the nominal interest rate and inflation are locally determinate.[5]

[5] See Woodford (2003), chapter 2, Proposition 2.8.

8 Case 2: Firms have limited attention, households have perfect information

In this section we solve the model under rational inattention on the side of firms, that is, we assume that decisionmakers in firms have limited attention and choose the allocation of their attention. For the moment, we continue to assume that households have perfect information to isolate the role of rational inattention by decisionmakers in firms.

8.1 Firms' attention problem

We now formalize the idea that decisionmakers cannot attend perfectly to all available information. Following Sims (2003), we model decisionmakers' limited attention as a constraint on information flow, and we let decisionmakers choose how to satisfy this constraint. In other words, decisionmakers decide what to focus on. For example, decisionmakers decide how to allocate their attention across their different decision problems. Furthermore, decisionmakers decide how to attend to the different factors that may affect an optimal decision. Formally, the attention problem of the decisionmaker in firm i reads

$$\max_{B(L),C(L)} E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t \frac{1}{2}\left(x_t - x_t^*\right)' H_x \left(x_t - x_t^*\right)\right], \tag{81}$$

subject to an equation linking an argument of the objective and a decision variable

$$\hat{p}_{it} - \hat{p}_{it}^* = p_{it} - p_{it}^*, \tag{82}$$

the equations characterizing the profit-maximizing decisions

$$p_{it}^* = \underbrace{A_1(L)\,\varepsilon_t^A}_{p_{it}^{A*}} + \underbrace{A_2(L)\,\varepsilon_t^R}_{p_{it}^{R*}} + \underbrace{A_3(L)\,\varepsilon_{it}^I}_{p_{it}^{I*}} \tag{83}$$

$$\hat{l}_{ijnt}^* = A_4(L)\,\varepsilon_{jnt}^{\chi}, \tag{84}$$

the equations specifying the actual decisions

$$p_{it} = \underbrace{B_1(L)\,\varepsilon_t^A + C_1(L)\,\nu_{it}^A}_{p_{it}^A} + \underbrace{B_2(L)\,\varepsilon_t^R + C_2(L)\,\nu_{it}^R}_{p_{it}^R} + \underbrace{B_3(L)\,\varepsilon_{it}^I + C_3(L)\,\nu_{it}^I}_{p_{it}^I} \tag{85}$$

$$\hat{l}_{ijnt} = B_4(L)\,\varepsilon_{jnt}^{\chi} + C_4(L)\,\nu_{ijnt}^{\chi}, \tag{86}$$

and the information flow constraint

$$\mathcal{I}\left(\left\{p_{it}^{A*},p_{it}^{R*},p_{it}^{I*},\hat{l}_{i11t}^{*},\ldots,\hat{l}_{iJ(N-1)t}^{*}\right\};\left\{p_{it}^{A},p_{it}^{R},p_{it}^{I},\hat{l}_{i11t},\ldots,\hat{l}_{iJ(N-1)t}\right\}\right)\leq\kappa. \tag{87}$$

Here ν_{it}^{A}, ν_{it}^{R}, ν_{it}^{I} and ν_{ijnt}^{χ} follow idiosyncratic Gaussian white noise processes with unit variance that are mutually independent and independent of all other shocks in the economy.

Expression (81) is minus the expected loss in profits due to suboptimal decisions. Equation (82) states that the mistake in the relative price of good i equals the mistake in the dollar price of good i. This equation is important because the objective depends on the mistake in the relative price of good i, while the decisionmaker sets the dollar price of good i. The equation follows from the fact that $\left(\hat{P}_{it}/\hat{P}_{it}^{*}\right)=\left(P_{it}/P_{it}^{*}\right)$. Equations (83)-(84) characterize the profit-maximizing decisions. Here $A_1(L)$, $A_2(L)$, $A_3(L)$ and $A_4(L)$ are infinite-order lag polynomials. These two equations follow from equations (46)-(47), $p_{it}=\hat{p}_{it}+p_t$ and the stochastic processes for p_t, c_t, a_t, a_{it} and $\tilde{w}_{jnt}$. Equations (85)-(86) specify the actual decisions. Choosing the lag polynomials $B_1(L)$ and $C_1(L)$ to $B_4(L)$ and $C_4(L)$ amounts to choosing the stochastic processes for the actual decisions. These lag polynomials imply a mapping between shocks and price setting and factor mix decisions. For example, if $B_1(L)=A_1(L)$ and $C_1(L)=0$, the price set by the decisionmaker responds perfectly to aggregate technology shocks. The information flow constraint (87) introduces limited attention on the side of the decisionmaker. Limited attention is modeled as a constraint on information flow. The operator $\mathcal{I}$ measures information flow between stochastic processes.[6] The left-hand side of (87) measures how much information the actual behavior contains about the profit-maximizing behavior. The parameter κ is the bound on information flow. The parameter κ indexes the decisionmaker's total attention devoted to price setting and factor mix decisions. We will choose the parameter κ such that the private marginal value of information flow is small.

Note that we assume that the noise shocks ν_{it}^{A}, ν_{it}^{R}, ν_{it}^{I} and ν_{ijnt}^{χ} follow Gaussian processes. It turns out that Gaussianity is optimal because the objective is quadratic and the profit-maximizing decisions follow Gaussian processes.[7] Furthermore, note that we assume that ν_{it}^{A}, ν_{it}^{R}, ν_{it}^{I} and ν_{ijnt}^{χ} are mutually independent. In the future, we also plan to

[6] For a definition of the operator $\mathcal{I}$, see equations (1)-(4) in Maćkowiak and Wiederholt (2008).

[7] See Sims (2006) or Maćkowiak and Wiederholt (2008).

study the case where these noise shocks can be correlated. Finally, note that we assume that these noise shocks are idiosyncratic. This assumption accords well with the idea that the friction is the decisionmaker's limited attention rather than the availability of information.

8.2 Computing the equilibrium

We use an iterative procedure to solve for the equilibrium of the model. First, we make a guess concerning the process for the profit-maximizing price and the process for the profit-maximizing factor mix. See equations (83)-(84). Second, we solve the firms' attention problem (81)-(87). Third, we aggregate the individual prices to obtain the aggregate price level:

$$p_t = \frac{1}{I}\sum_{i=1}^{I} p_{it}. \tag{88}$$

Fourth, we compute the aggregate dynamics implied by those price level dynamics. The following equations have to be satisfied in equilibrium:

$$r_t = \rho_R r_{t-1} + (1-\rho_R)\left[\phi_\pi\left(p_t - p_{t-1}\right) + \phi_y y_t\right] + \varepsilon_t^R, \tag{89}$$

$$c_t = E_t\left[-\frac{1}{\gamma}\left(r_t - p_{t+1} + p_t\right) + c_{t+1}\right], \tag{90}$$

$$y_t = c_t, \tag{91}$$

$$y_t = a_t + \alpha l_t, \tag{92}$$

$$a_t = \rho_A a_{t-1} + \varepsilon_t^A, \tag{93}$$

$$\tilde{w}_t = \gamma c_t + \psi l_t. \tag{94}$$

The first equation is the monetary policy rule. The second equation follows from optimal consumption behavior by households. The third equation follows from the requirement that output has to equal demand. The fourth equation follows from the production function. The fifth equation is the assumed process for aggregate technology. The sixth equation follows from optimal wage setting by households. We employ a standard solution method for linear rational expectations models to solve the system of equations containing the price level dynamics and those six equations. We obtain the law of motion for $(r_t, c_t, y_t, l_t, a_t, \tilde{w}_t)$

implied by the price level dynamics. Fifth, we compute the law of motion for the profit-maximizing price from equation (46) which we reproduce for convenience:

$$p_{it}^* = p_t + \frac{\frac{1-\alpha}{\alpha}}{1+\frac{1-\alpha}{\alpha}\vartheta} c_t - \frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\vartheta} (a_t + a_{it}) + \frac{1}{1+\frac{1-\alpha}{\alpha}\vartheta} \tilde{w}_t.$$

Optimal consumption behavior by households implies that $\vartheta = \theta$. If the process for the profit-maximizing price differs from our guess, we update our guess until we reach a fixed point.

Finally, we compute the equilibrium factor mix and the equilibrium relative wage rates. This is explained in Appendix D.

8.3 Benchmark parameter values and solution

In this section we report the numerical solution of the model for the following parameter values. We set $\beta = 0.99$, $\gamma = 1$, $\psi = 1$, $\theta = 4$, $\alpha = 2/3$ and $\eta = 4$.

To set the parameters governing the process for aggregate technology, equation (93), we consider quarterly U.S. data from 1960 Q1 to 2006 Q4. We first compute a time series for aggregate technology, a_t, using equation (92) and measures of y_t and l_t. We use the log of real output per person, detrended with a linear trend, as a measure of y_t. We use the log of hours worked per person, demeaned, as a measure of l_t.[8] We then fit equation (93) to the time series for a_t obtaining $\rho_A = 0.96$ and a standard deviation of the innovation equal to 0.0085. In the benchmark economy we set $\rho_A = 0.95$ and we set the standard deviation of ε_t^A equal to 0.0085.

To set the parameters of the Taylor rule we consider quarterly U.S. data on the Federal Funds rate, inflation and real GDP from 1960 Q1 to 2006 Q4.[9] We fit the Taylor rule (89) to the data obtaining $\rho_R = 0.89$, $\phi_\pi = 1.53$, $\phi_y = 0.33$, and a standard deviation of the innovation equal to 0.0021. In the benchmark economy we set $\rho_R = 0.9$, $\phi_\pi = 1.5$, $\phi_y = 0.33$, and the standard deviation of ε_t^R equal to 0.0021.

[8] We use data for the non-farm business sector. The source of the data is the website of the Federal Reserve Bank of St.Louis.

[9] We compute a time series for four-quarter inflation rate from the price index for personal consumption expenditures excluding food and energy. We compute a time series for percentage deviations of real GDP from potential real GDP. The sources of the data are the websites of the Federal Reserve Bank of St.Louis and the Congressional Budget Office.

We assume that firm-specific productivity follows a first-order autoregressive process. Recent papers calibrate the autocorrelation of firm-specific productivity to be about two-thirds in monthly data, e.g., Klenow and Willis (2007) use 0.68, Midrigan (2006) uses 0.5, and Nakamura and Steinsson (2007) use 0.66. Since $(2/3)^3$ equals about 0.3, we set the autocorrelation of firm-specific productivity in our quarterly model equal to 0.3. We then choose the standard deviation of the innovation to firm-specific productivity such that the average absolute size of price changes in our model equals 9.7 percent under perfect information. The value 9.7 percent is the average absolute size of price changes excluding sales reported in Klenow and Kryvtsov (2007). This yields a standard deviation of the innovation to firm-specific productivity equal to 0.22.

We assume that labor-specific preference shocks follow a white noise process. This simplifying assumption implies that solving for relative wage rates and relative labor inputs is straightforward. See Appendix D. To set the standard deviation of labor-specific preference shocks we proceed as follows. Autor, Katz and Kearney (2005) report the variance of log hourly wages of men in the U.S. between 1975 and 2003. The average variance of log hourly wages of men in this period was 0.32. We choose the variance of χ_{jnt} such that the variance of $\tilde{w}_{jnt}$ in our model equals 0.32 under perfect information. This yields a standard deviation of labor-specific preference shocks equal to 2.83. See equation (79) and recall that $\psi = 1$ and $\eta = 4$. We set the number of different types of labor that a firm hires to $JN = 100$.

We compute the solution of the model by fixing the marginal value of information flow. The total information flow, κ, is then determined within the model. The idea is the following. When the marginal value of information flow is high, decisionmakers have a high incentive to increase information flow in order to take better decision. In contrast, when the marginal value of information flow is low, decisionmakers have little incentive to increase information flow. We set the marginal value of information flow equal to 0.25 percent of a firm's steady state output.

We first report the optimal allocation of attention at the rational inattention fixed point. The total information flow at the solution equals 133 bits. The decisionmaker in a firm allocates: 2.46 bits of information flow (his/her attention) to tracking firm-specific productivity, 1.31 bits of information flow to tracking each relative wage rate, 0.76 bits of information flow to tracking aggregate technology, and 0.41 bits of information flow to tracking monetary policy. The expected per period loss in profits due to imperfect tracking of firm-specific productivity equals 0.18 percent of the firm's steady state output. The expected per period loss in profits due to imperfect tracking of aggregate technology equals 0.12 percent of the firm's steady state output. The expected per period loss in profits due to imperfect tracking of monetary policy equals 0.07 percent of the firm's steady state output. Together these numbers imply that the expected per period loss in profits due to deviations of the actual price from the profit-maximizing price equals 0.37 percent of the firm's steady state output. We think this is a reasonable number.

Figures 1 and 2 show impulse responses of the price level, inflation, output, and the nominal interest rate at the rational inattention fixed point (green lines with circles). For comparison the figures also include impulse responses of the same variables at the perfect information equilibrium (blue lines with points). All impulse responses are to shocks of one standard deviation. All impulse responses are drawn such that an impulse response equal to one means "a one percent deviation from the non-stochastic steady state". Time is measured in quarters along horizontal axes.

Consider Figure 1. The price level shows a dampened and delayed response to a monetary policy shock compared with the case of perfect information. The response of inflation to a monetary policy shock is persistent. Output falls after a positive innovation in the Taylor rule and the decline in output is persistent. The nominal interest rate increases on impact and converges slowly to zero. The impulse responses to a monetary policy shock under rational inattention differ markedly from the impulse responses to a monetary policy shock under perfect information. Under perfect information the price level adjusts fully on impact to a monetary policy shock, there are no real effects, and the nominal interest rate fails to change.

Consider Figure 2. The price level and inflation show a dampened response to an aggregate technology shock compared with the case of perfect information. The output gap is negative for a few quarters after the shock. Output and the nominal interest rate show hump-shaped impulse responses to an aggregate technology shock. Note that under rational inattention the response of the price level to an aggregate technology shock is less dampened and less delayed than the response of the price level to a monetary policy shock. The reason is the optimal allocation of attention. Since decisionmakers in firms allocate about twice as much attention to aggregate technology than to monetary policy, prices respond faster to aggregate technology shocks than to monetary policy shocks. Therefore, the output gap is negative for only 5 quarters after an aggregate technology shock, while the output gap is negative for more than 10 quarters after a monetary policy shock.[10]

Figure 3 shows the impulse response of an individual price to a firm-specific productivity shock. Note that prices respond almost perfectly to firm-specific productivity shocks. The reason is the optimal allocation of attention.

8.4 Comparison to the Calvo model

For comparison, we solved the Calvo model for the same parameter values and assuming that prices change every 2.5 quarters on average.[11] Figures 4 and 5 show the impulse responses in the benchmark economy with rational inattention (green lines with circles) and the impulse responses in the perfect information, Calvo model (red lines with crosses). The impulse responses to a monetary policy shock are very similar in the two models. In contrast, the impulse responses to an aggregate technology shock are quite different in the two models. Inflation responds to a monetary policy shock by the same amount on impact in the benchmark economy and in the Calvo model, while inflation responds to an aggregate technology shock twice more strongly on impact in the benchmark economy than in the Calvo model. The reason is that decisionmakers in firms in the benchmark economy allocate about twice as much attention to aggregate technology than to monetary policy.

Firms in the benchmark economy and firms in the Calvo model experience profit losses due to deviations of the actual price from the profit-maximizing price. It turns out that

[10] See also Paciello (2007). Paciello solves the white noise case of a similar model analytically, where white noise case means that: (i) all exogenous processes are white noise processes, (ii) there is no lagged interest rate in the Taylor rule, and (iii) the price level instead of inflation appears in the Taylor rule. The analytical solution in the white noise case helps to understand in more detail the differential response of prices to aggregate technology shocks and to monetary policy shocks.

[11] Klenow and Kryvtsov (2007) find that the median price duration excluding sales is 7.2 months, or about 2.5 quarters.

profit losses due to deviations of the actual price from the profit-maximizing price are an order of magnitude smaller than in the Calvo model generating the same real effects. Specifically, the expected loss in profits due to suboptimal price responses to aggregate conditions is 12 times smaller than in the Calvo model; and the expected loss in profits due to suboptimal price responses to firm-specific conditions is 25 times smaller than in the Calvo model. One reason is that in the benchmark economy prices respond slowly to monetary policy shocks, faster to aggregate technology shocks, and very fast to micro-level shocks. In contrast, in the Calvo model prices respond slowly to all those shocks. Another reason is that under rational inattention deviations of the actual price from the profit-maximizing price are less likely to be extreme than in the Calvo model.

8.5 Varying parameter values

Figure 6 compares the benchmark economy to an economy with a higher degree of real rigidity.[12] We set $\psi = 0$ implying that the coefficient on aggregate output in the equation for the profit-maximizing price falls from 1 to 0.5 (after substituting in the wage equation). Real effects of monetary policy shocks become larger and more persistent. After a monetary policy shock the output gap is now negative for about 20 quarters instead of 10 quarters. At the same time, profit losses due to imperfect tracking of aggregate conditions decrease. The expected loss in profits due to imperfect tracking of aggregate technology falls by 25 percent; and the expected loss in profits due to imperfect tracking of monetary policy falls by 60 percent.

Figure 7 compares the benchmark economy to an economy with larger monetary policy shocks. We increase the standard deviation of monetary policy shocks by roughly a factor of two, from 0.0021 to 0.004. This matches the standard deviation of monetary policy shocks estimated by Justiniano and Primiceri (2006) for the high inflation episode in the 1970s.[13] At the rational inattention fixed point with larger monetary policy shocks firms allocate 0.84 bits to monetary policy, an increase by 100 percent compared to the benchmark

[12] Ball and Romer (1990) refer to the elasticity of the profit-maximizing price with respect to aggregate output as the degree of real rigidity. A low elasticity corresponds to a high degree of real rigidity.

[13] Justiniano and Primiceri (2006) estimate a DSGE model that allows for time variation in the size of shocks.

economy.[14] Since firms allocate more attention to monetary policy, a monetary policy shock of a given size has smaller real effects and larger inflationary effects.[15] However, the additional attention that decisionmakers pay to monetary policy is not sufficient to compensate fully for the fact that the average size of monetary policy shocks has increased. The firms' tracking problem has become more complicated. Therefore, the expected loss in profits due to imperfect tracking of monetary policy almost doubles and the variance of output due to monetary policy shocks increases.

9 Case 3: Firms and households have limited attention

In this section we study the implications of adding rational inattention by households. To get a first idea of how rational inattention by households affects the equilibrium, we make three simplifying assumptions. First, we study the optimal allocation of attention by an individual household assuming that all other households have perfect information and firms have limited attention, i.e., we study the optimal allocation of attention by an individual household at the fixed point derived in Section 8. Second, we assume that the household sets a vector of real wage rates instead of a vector of nominal wage rates. Third, we assume that the household has linear disutility of labor, i.e., $\psi = 0$. One can show analytically that the last two assumptions imply that the optimal wage setting behavior under perfect information and under limited attention is given by

$$\tilde{w}_{jt} = \gamma c_{jt}.$$

Then the household is always on his labor supply curve because he only needs to know his own consumption to be on the labor supply curve. This allows us to study in isolation the implications of rational inattention by households for consumption behavior.

[14] We are holding constant the marginal value of information flow. The total information flow is then determined within the model.

[15] In Figure 7 one must divide an impulse response in the economy with larger monetary policy shocks by roughly one-half to obtain an impulse response to a shock of the same size as in the benchmark economy.

9.1 Households' attention problem

The attention problem of household j reads

$$\max_{B(L),C(L)} \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[\frac{1}{2} (x_t - x_t^*)' H_0 (x_t - x_t^*) + (x_t - x_t^*)' H_1 \left(x_{t+1} - x_{t+1}^* \right) \right], \tag{95}$$

subject to an equation linking an argument of the objective and $1+N$ decision variables

$$\tilde{b}_{jt} - \tilde{b}_{jt}^* = -\frac{1}{\omega_B} \sum_{s=0}^{t} \left(\frac{1}{\beta} \right)^{t-s} \left[\left(c_{js} - c_{js}^* \right) + \frac{\omega_W (\zeta - 1)}{N} \sum_{n=1}^{N} \left(\tilde{w}_{jns} - \tilde{w}_{jns}^* \right) \right], \tag{96}$$

the equations characterizing the optimal behavior under perfect information

$$c_{jt}^* = \underbrace{A_1(L)\,\varepsilon_t^A}_{c_{jt}^{A*}} + \underbrace{A_2(L)\,\varepsilon_t^R}_{c_{jt}^{R*}} \tag{97}$$

$$\tilde{w}_{jnt}^* = \gamma c_{jt}^* + \underbrace{A_3(L)\,\varepsilon_{jnt}^{\chi}}_{\tilde{w}_{jnt}^{\chi*}} \tag{98}$$

$$\hat{c}_{ijt}^* = A_4(L)\,\varepsilon_{it}^I, \tag{99}$$

the equations specifying the actual behavior

$$c_{jt} = \underbrace{B_1(L)\,\varepsilon_t^A + C_1(L)\,\nu_{jt}^A}_{c_{jt}^A} + \underbrace{B_2(L)\,\varepsilon_t^R + C_2(L)\,\nu_{jt}^R}_{c_{jt}^R} \tag{100}$$

$$\tilde{w}_{jnt} = \gamma c_{jt} + \underbrace{B_3(L)\,\varepsilon_{jnt}^{\chi} + C_3(L)\,\nu_{jnt}^{\chi}}_{\tilde{w}_{jnt}^{\chi}} \tag{101}$$

$$\hat{c}_{ijt} = B_4(L)\,\varepsilon_{it}^I + C_4(L)\,\nu_{ijt}^I, \tag{102}$$

and the information flow constraint

$$\mathcal{I}\left(\left\{ c_{jt}^{A*}, c_{jt}^{R*}, \tilde{w}_{j1t}^{\chi*}, \ldots, \tilde{w}_{jNt}^{\chi*}, \hat{c}_{1jt}^*, \ldots, \hat{c}_{I-1jt}^* \right\} ; \left\{ c_{jt}^A, c_{jt}^R, \tilde{w}_{j1t}^{\chi}, \ldots, \tilde{w}_{jNt}^{\chi}, \hat{c}_{1jt}, \ldots, \hat{c}_{I-1jt} \right\} \right) \leq \kappa. \tag{103}$$

Here ν_{jt}^A, ν_{jt}^R, ν_{jnt}^{χ} and ν_{ijt}^I follow idiosyncratic Gaussian white noise processes with unit variance that are mutually independent and independent of all other shocks in the economy. All the $A(L)$, $B(L)$ and $C(L)$ are infinite-order lag polynomials.

Note that we assume that the household chooses a consumption vector and a vector of real wage rates. Bond holdings then follow from equation (96).

Furthermore, we assume that in period $t = -1$ the economy is in the non-stochastic steady state and the household knows that the economy is in the non-stochastic steady state.

9.2 Benchmark parameter values and solution

We assume the same parameter values as in the benchmark economy in Section 8.3 apart from $\psi = 0$. We have to choose values for some additional parameters: ω_B and ω_W as well as the household's marginal value of information flow. We set $\omega_B = \omega_W = 1$. We set the household's marginal value of information flow equal to 0.02 percent of the household's steady state composite consumption.

The optimal allocation of attention by the household has the following features. The household allocates 0.83 bits of information flow to tracking monetary policy and 0.74 bits of information flow to tracking aggregate technology. The expected per period loss in utility due to imperfect tracking of monetary policy equals 0.02 percent of the household's steady state composite consumption. The expected per period loss in utility due to imperfect tracking of aggregate technology also equals 0.02 percent of the household's steady state composite consumption. Figure 8 shows the impulse response of composite consumption by the individual household to a monetary policy shock (upper panel) and to an aggregate technology shock (lower panel). In each panel, the blue line with points is the impulse response under perfect information, while the green line with circles is the impulse response under limited attention. We would like to point out four results. First, there are sizeable differences between the impulse responses for consumption under perfect information and the impulse responses for consumption under rational inattention despite the fact that the utility loss from deviations from the perfect information behavior is very small and the marginal value of information flow is very low. Second, the impulse response of consumption to a monetary policy shock under rational inattention is hump-shaped, while the impulse response under perfect information is monotonic. Third, consumption under rational inattention differs from consumption under perfect information, but in the long run the difference between consumption under rational inattention and consumption under perfect information goes to zero. Similarly, we find that bond holdings under rational inattention differ from bond holdings under perfect information, but in the long run the difference between bond holdings under rational inattention and bond holdings under perfect information goes to zero. Fourth, the impulse responses of consumption under rational inattention look similar to the impulse responses of consumption in a model with habit formation.

10 Conclusion

We have studied a dynamic stochastic general equilibrium model with rational inattention. The impulse responses of prices under rational inattention by firms have several properties of empirical impulse response functions, e.g., (i) prices respond slowly to monetary policy shocks, (ii) prices respond faster to aggregate technology shocks, and (iii) prices respond very fast to disaggregate shocks.[16] Furthermore, profit losses due to deviations of the actual price from the profit-maximizing price are an order of magnitude smaller than in the Calvo model that generates the same real effects.

In addition, we have presented some results concerning the implications of adding rational inattention by households. The impulse responses of consumption under rational inattention by households look similar to the impulse responses of consumption in a model with habit formation.

[16] For empirical evidence on the response of prices to disaggregate shocks, see Boivin, Giannoni, and Mihov (2007).

A Proof of Proposition 1

First, let x_t denote the vector of date t variables appearing in objective (39) that the firm can affect

$$x_t = \begin{pmatrix} \hat{p}_{it} & \hat{l}_{i11t} & \cdots & \hat{l}_{iJ(N-1)t} \end{pmatrix}'.$$

Let z_t denote the vector of date t variables appearing in objective (39) that the firm takes as given

$$z_t = \begin{pmatrix} c_{1t} & \cdots & c_{Jt} & c_t & a_t & a_{it} & \tilde{w}_{11t} & \cdots & \tilde{w}_{JNt} \end{pmatrix}'.$$

Let f denote the functional inside the expectation operator in objective (39). Let $\tilde{f}$ denote the second-order Taylor approximation of f at the non-stochastic steady state. We have

$$\begin{aligned} & E_{i,-1}\left[\tilde{f}\left(x_0, z_0, x_1, z_1, x_2, z_2, \ldots\right)\right] \\ = \; & E_{i,-1}\left[\begin{array}{c} f\left(0,0,0,0,0,0,\ldots\right) \\ +\sum_{t=0}^{\infty}\beta^t\left(h_x' x_t + h_z' z_t + \frac{1}{2}x_t' H_x x_t + x_t' H_{xz} z_t + \frac{1}{2} z_t' H_z z_t\right) \end{array}\right], \end{aligned} \tag{104}$$

where h_x is $\left(1/\beta^t\right)$ times the $(JN \times 1)$ vector of first derivatives of f with respect to x_t evaluated at the non-stochastic steady state, h_z is $\left(1/\beta^t\right)$ times the $(J+3+JN \times 1)$ vector of first derivatives of f with respect to z_t evaluated at the non-stochastic steady state, H_x is $\left(1/\beta^t\right)$ times the $(JN \times JN)$ matrix of second derivatives of f with respect to x_t evaluated at the non-stochastic steady state, H_z is $\left(1/\beta^t\right)$ times the $(J+3+JN \times J+3+JN)$ matrix of second derivatives of f with respect to z_t evaluated at the non-stochastic steady state, and H_{xz} is $\left(1/\beta^t\right)$ times the $(JN \times J+3+JN)$ matrix of cross derivatives. Second, in this appendix let y_t denote the following vector

$$y_t = \begin{pmatrix} x_t \\ z_t \\ 1 \end{pmatrix}.$$

Conditions (42)-(43) imply that for all m and n

$$\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left|y_{mt} y_{nt}\right| < \infty. \tag{105}$$

It follows that

$$E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t y_{mt}y_{nt}\right]=\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[y_{mt}y_{nt}\right]. \tag{106}$$

See Rao (1973), p. 111. Furthermore, conditions (42)-(43) also imply that for all m and n the sequence $\left\{\sum_{t=0}^{T}\beta^t E_{i,-1}\left[y_{mt}y_{nt}\right]\right\}_{T=0}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$. It follows that

$$\begin{aligned}
& E_{i,-1}\left[\tilde{f}\left(x_0,z_0,x_1,z_1,x_2,z_2,\ldots\right)\right] \\
= \; & f\left(0,0,0,0,0,0,\ldots\right) \\
& +\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[h_x'x_t+h_z'z_t+\frac{1}{2}x_t'H_x x_t+x_t'H_{xz}z_t+\frac{1}{2}z_t'H_z z_t\right],
\end{aligned} \tag{107}$$

and both terms on the right-hand side of the last equation converges to an element in $\mathbb{R}$. Third, let x_t^* denote the vector x_t that satisfies

$$h_x+H_x x_t^*+H_{xz}z_t=0. \tag{108}$$

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

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

It follows that x_t^* satisfies conditions (42)-(43), implying that equation (107) holds for the vector x_t^*. Fourth,

$$\begin{aligned}
& E_{i,-1}\left[\tilde{f}\left(x_0,z_0,x_1,z_1,x_2,z_2,\ldots\right)\right]-E_{i,-1}\left[\tilde{f}\left(x_0^*,z_0,x_1^*,z_1,x_2^*,z_2,\ldots\right)\right] \\
= \; & \sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[h_x'\left(x_t-x_t^*\right)+\frac{1}{2}x_t'H_x x_t-\frac{1}{2}x_t^{*\prime}H_x x_t^*+\left(x_t-x_t^*\right)'H_{xz}z_t\right] \\
= \; & \sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}x_t'H_x x_t-\frac{1}{2}x_t^{*\prime}H_x x_t^*-\left(x_t-x_t^*\right)'H_x x_t^*\right] \\
= \; & \sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}\left(x_t-x_t^*\right)'H_x\left(x_t-x_t^*\right)\right],
\end{aligned} \tag{109}$$

where the first line follows from equation (107) and the second line follows from using equation (108) to substitute for $H_{xz}z_t$. Fifth, using the fact that $(1+\tau_p)=[\vartheta/(\vartheta-1)]$ and using the fact that the demand function (29) has the property $\varsigma\hat{P}_i^{-\vartheta}C=\hat{P}_i^{-\theta}C$ yields

$$h_x=0, \tag{110}$$

$$H_x = C_j^{-\gamma}\tilde{W}L_i \begin{bmatrix} -\frac{\vartheta}{\alpha}\left(1+\frac{1-\alpha}{\alpha}\vartheta\right) & 0 & \cdots & \cdots & 0 \\ 0 & -\frac{2}{\eta JN} & -\frac{1}{\eta JN} & \cdots & -\frac{1}{\eta JN} \\ \vdots & -\frac{1}{\eta JN} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -\frac{1}{\eta JN} \\ 0 & -\frac{1}{\eta JN} & \cdots & -\frac{1}{\eta JN} & -\frac{2}{\eta JN} \end{bmatrix}, \tag{111}$$

and

$$H_{xz} = C_j^{-\gamma}\tilde{W}L_i \begin{bmatrix} 0 & \cdots & 0 & \frac{\vartheta}{\alpha}\frac{1-\alpha}{\alpha} & -\frac{\vartheta}{\alpha}\frac{1}{\alpha} & -\frac{\vartheta}{\alpha}\frac{1}{\alpha} & \frac{\vartheta}{\alpha}\frac{1}{JN} & \cdots & \cdots & \frac{\vartheta}{\alpha}\frac{1}{JN} & \frac{\vartheta}{\alpha}\frac{1}{JN} \\ 0 & \cdots & 0 & 0 & 0 & 0 & -\frac{1}{JN} & 0 & \cdots & 0 & \frac{1}{JN} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -\frac{1}{JN} & \frac{1}{JN} \end{bmatrix}. \tag{112}$$

Substituting (110)-(112) into equation (108) yields that x_t^* is given by

$$\hat{p}_{it}^* = \frac{\frac{1-\alpha}{\alpha}}{1+\frac{1-\alpha}{\alpha}\vartheta}c_t - \frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\vartheta}\left(a_t + a_{it}\right) + \frac{1}{1+\frac{1-\alpha}{\alpha}\vartheta}\left(\frac{1}{JN}\sum_{j=1}^{J}\sum_{n=1}^{N}\tilde{w}_{jnt}\right), \tag{113}$$

and

$$\hat{l}_{ijnt}^* + \sum_{jn\neq JN}\hat{l}_{ijnt}^* = -\eta\left(\tilde{w}_{jnt} - \tilde{w}_{JNt}\right). \tag{114}$$

Summing equation (114) over all $jn \neq JN$ yields

$$\sum_{jn\neq JN}\hat{l}_{ijnt}^* = -\eta\frac{1}{JN}\sum_{j=1}^{J}\sum_{n=1}^{N}\tilde{w}_{jnt} + \eta\tilde{w}_{JNt}.$$

Substituting the last equation into equation (114) yields

$$\hat{l}_{ijnt}^* = -\eta\left(\tilde{w}_{jnt} - \frac{1}{JN}\sum_{j=1}^{J}\sum_{n=1}^{N}\tilde{w}_{jnt}\right). \tag{115}$$

B Proof of Proposition 2

Same steps as in Appendix A.

C Solution under perfect information

First, $y_{it} = c_{it} = (1/J)\sum_{j=1}^{J} c_{ijt}$, equation $c_{ijt} = \hat{c}_{ijt} + c_{jt}$ and equations (65), (66), (67) and (69) imply that

$$y_t = c_t.$$

Second, equations (66), (68), (71) and (11) imply that

$$y_t = a_t + \alpha l_t.$$

Third, equations (46), (69), (70) and (11) imply that

$$0 = \frac{1-\alpha}{\alpha} c_t - \frac{1}{\alpha} a_t + \tilde{w}_t.$$

Fourth, equations (64), $\tilde{w}_{jnt} = w_{jnt} - p_t$, $\tilde{w}_t = w_t - p_t$, (70), (67) and (12) imply that

$$\tilde{w}_t = \gamma c_t + \psi l_t.$$

Solving the last four equations for the endogenous variables y_t, c_t, l_t and $\tilde{w}_t$ yields (73)-(75). Fifth, equations (63) and (67) imply

$$c_t = E_t \left[-\frac{1}{\gamma} \left(r_t - \pi_{t+1} \right) + c_{t+1} \right].$$

Substituting the solution for c_t into the last equation yields (76).

D Solving for equilibrium relative wage rates

Solving for equilibrium relative wage rates is not easy, because when households set wage rates they take into account that limited attention by decisionmakers in firms lowers the wage elasticity of labor demand. To make the derivation as clear as possible, we assume that all labor-specific preference shocks, χ_{jnt}, follow a Gaussian white noise process with variance σ_χ^2.

First, let $\hat{l}_{ijnt} = l_{ijnt} - l_{it}$ denote firm i's relative input of type jn labor and let $\hat{w}_{jnt} = \tilde{w}_{jnt} - \tilde{w}_t$ denote the relative wage rate for type jn labor. The profit-maximizing factor mix is given by

$$\hat{l}_{ijnt}^* = -\eta \hat{w}_{jnt}. \tag{116}$$

See Proposition 1. We guess that in equilibrium

$$\hat{w}_{jnt} = A\chi_{jnt}, \tag{117}$$

where A is an undetermined coefficient. The assumption that χ_{jnt} follows a Gaussian white noise process implies that both $\hat{w}_{jnt}$ and $\hat{l}^*_{ijnt}$ follow Gaussian white noise processes.

Second, devoting an information flow of κ_χ to tracking a variable that follows a Gaussian white noise process so as to minimize the mean squared error yields the following behavior under rational inattention

$$\hat{l}^{RI}_{ijnt} = \left(1 - \frac{1}{2^{2\kappa_\chi}}\right)\hat{l}^*_{ijnt} + \sqrt{\frac{1}{2^{2\kappa_\chi}} - \frac{1}{2^{4\kappa_\chi}}}\sqrt{Var\left(\hat{l}^*_{ijnt}\right)}\nu^\chi_{ijnt}, \tag{118}$$

where ν^χ_{ijnt} follows an independent Gaussian white noise process with unit variance. See, for example, Proposition 3 in Maćkowiak and Wiederholt (2008). Using (116) to substitute for $\hat{l}^*_{ijnt}$ in (118) yields

$$\hat{l}^{RI}_{ijnt} = -\eta\left(1 - \frac{1}{2^{2\kappa_\chi}}\right)\left(\hat{w}_{jnt} - \sqrt{\frac{1}{2^{2\kappa_\chi} - 1}}\sqrt{Var\left(\hat{w}_{jnt}\right)}\nu^\chi_{ijnt}\right). \tag{119}$$

It is now easy to verify that the signal

$$s_{ijnt} = \hat{w}_{jnt} - \sqrt{\frac{1}{2^{2\kappa_\chi} - 1}}\sqrt{Var\left(\hat{w}_{jnt}\right)}\nu^\chi_{ijnt}, \tag{120}$$

has the property

$$\hat{l}^{RI}_{ijnt} = E\left[\hat{l}^*_{ijnt}|s_{ijnt}, s_{ijnt-1}, \ldots\right].$$

Therefore, one can interpret the behavior under rational inattention as being due to the fact that decisionmakers in firms pay limited attention to the relative wage rate for type jn labor. Furthermore, note that limited attention by decisionmakers in firms lowers the wage elasticity of labor demand from η to $\eta\left(1 - \frac{1}{2^{2\kappa_\chi}}\right)$.

Third, let $l_{jnt} = \frac{1}{I}\sum_{i=1}^{I} l_{ijnt}$ and $l_t = \frac{1}{I}\sum_{i=1}^{I} l_{it}$. Computing the average of (119) over all i and using the fact that noise is idiosyncratic yields

$$l_{jnt} - l_t = -\eta\left(1 - \frac{1}{2^{2\kappa_\chi}}\right)\hat{w}_{jnt}. \tag{121}$$

Exponentiating both sides of (121), multiplying by L_{jn} and using the fact that in the non-stochastic steady state $L_{jn} = \left(\frac{W_{jn}}{W}\right)^{-\eta} L$ yields

$$L_{jnt} = \left(\frac{W_{jn}}{W}\right)^{-\frac{\eta}{2^{2\kappa_\chi}}}\left(\frac{W_{jnt}}{W_t}\right)^{-\eta\left(1 - \frac{1}{2^{2\kappa_\chi}}\right)} L_t. \tag{122}$$

Note that this demand function for type jn labor has the form (48)-(50) so long as $\eta\left(1-\frac{1}{2^{2\kappa_{\chi}}}\right)>1$.

Fourth, since households have perfect information, household j then sets a nominal wage rate for type jn labor that satisfies

$$\tilde{w}_{jnt}=\frac{\gamma}{1+\psi\zeta}c_{jt}+\frac{\psi}{1+\psi\zeta}\left(\zeta\tilde{w}_{t}+l_{t}\right)+\frac{1}{1+\psi\zeta}\chi_{jnt}, \tag{123}$$

with $\zeta=\eta\left(1-\frac{1}{2^{2\kappa_{\chi}}}\right)$. See Proposition 2. Computing the average of (123) over all jn and using (12), (67) and (70) yields

$$\tilde{w}_{t}=\gamma c_{t}+\psi l_{t}. \tag{124}$$

Substituting (124) into (123) and using $c_{jt}=c_{t}$ yields

$$\tilde{w}_{jnt}-\tilde{w}_{t}=\frac{1}{1+\psi\zeta}\chi_{jnt}, \tag{125}$$

where $\zeta=\eta\left(1-\frac{1}{2^{2\kappa_{\chi}}}\right)$. Comparing (125) to the guess (117) shows that the guess (117) is correct with

$$A=\frac{1}{1+\psi\eta\left(1-\frac{1}{2^{2\kappa_{\chi}}}\right)}. \tag{126}$$

Finally, we still need to solve for the equilibrium attention that decisionmakers in firms allocate to the relative wage rate for type jn labor: κ_{χ}. Equations (116), (117), (119) and (126) imply that the mean squared error in the relative input of type jn labor equals

$$E\left[\left(\hat{l}_{ijnt}^{*}-\hat{l}_{ijnt}^{RI}\right)^{2}\right]=\frac{1}{2^{2\kappa_{\chi}}}\frac{\eta^{2}}{\left[1+\psi\eta\left(1-\frac{1}{2^{2\kappa_{\chi}}}\right)\right]^{2}}\sigma_{\chi}^{2}.$$

The derivative of the mean squared error with respect to κ_{χ} equals

$$\frac{\partial E\left[\left(\hat{l}_{ijnt}^{*}-\hat{l}_{ijnt}^{RI}\right)^{2}\right]}{\partial\kappa_{\chi}}=-2\ln\left(2\right)\frac{1}{2^{2\kappa_{\chi}}}\frac{1+\psi\eta\left(1+\frac{1}{2^{2\kappa_{\chi}}}\right)}{\left[1+\psi\eta\left(1-\frac{1}{2^{2\kappa_{\chi}}}\right)\right]^{3}}\eta^{2}\sigma_{\chi}^{2}.$$

It follows from objective (81) that the marginal value of paying attention to the relative wage rate for type jn labor equals

$$\lambda_{\chi}=\frac{1}{1-\beta}\frac{1}{\eta JN}2\ln\left(2\right)\frac{1}{2^{2\kappa_{\chi}}}\frac{1+\psi\eta\left(1+\frac{1}{2^{2\kappa_{\chi}}}\right)}{\left[1+\psi\eta\left(1-\frac{1}{2^{2\kappa_{\chi}}}\right)\right]^{3}}\eta^{2}\sigma_{\chi}^{2}.$$

By equating λ_{χ} to our value for the marginal value of information flow we obtain κ_{χ}.

E Solving the Calvo model

If we assume that firms and households have perfect information but firms face a Calvo friction, we obtain the following version of the New Keynesian Phillips curve

$$\pi_t = \frac{(1-\lambda)(1-\lambda\beta)}{\lambda} \frac{\frac{\psi}{\alpha} + \gamma + \frac{1-\alpha}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\theta} \left(c_t - c_t^f\right) + \beta E_t\left[\pi_{t+1}\right], \tag{127}$$

where $(1-\lambda)$ is the fraction of goods prices that change every period and c_t^f is the flexible price solution given by equation (73). The aggregate dynamics are obtained by solving the system containing equations (89)-(94) and equation (127). The solution of the Calvo model reported in Figures 4-5 sets $\lambda = 0.6$.

References

[1] Autor, David H., Lawrence F. Katz and Melissa S. Kearney (2005): "Rising Wage Inequality: the Role of Composition and Prices." NBER Working Paper 11628.

[2] Ball, Lawrence and David Romer (1990): "Real Rigidities and the Non-Neutrality of Money." Review of Economic Studies, 57(2): 183-203.

[3] Boivin, Jean, Marc Giannoni and Ilian Mihov (2007): "Sticky Prices and Monetary Policy: Evidence from Disaggregated U.S. Data." Discussion paper, HEC Montreal, Columbia University and INSEAD.

[4] Justiniano, Alejandro and Giorgio E. Primiceri (2006): "The Time Varying Volatility of Macroeconomic Fluctuations." NBER Working Paper 12022.

[5] Klenow, Peter J. and Oleksiy Kryvtsov (2007): "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" Discussion paper, Stanford University and Bank of Canada.

[6] Klenow, Peter J. and Jonathan L. Willis (2007): "Sticky Information and Sticky Prices." Journal of Monetary Economics, 54 (sup. 1), 79-99 .

[7] Lucas, Robert E. Jr. (1972): "Expectations and the Neutrality of Money." Journal of Economic Theory, 4(2): 103-124.

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

[9] Lorenzoni, Guido (2008): "A Theory of Demand Shocks." Discussion paper, MIT and Federal Reserve Bank of Chicago.

[10] Maćkowiak, Bartosz and Mirko Wiederholt (2008): "Optimal Sticky Prices under Rational Inattention." forthcoming, American Economic Review.

[11] 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.

[12] Midrigan, Virgiliu (2006): "Menu Costs, Multi-Product Firms and Aggregate Fluctuations." Discussion paper, Ohio State University.

[13] Nakamura, Emi and Jón Steinsson (2007): "Five Facts About Prices: A Reevaluation of Menu Cost Models." Discussion paper, Harvard University.

[14] Paciello, Luigi (2007): "The Response of Prices to Technology and Monetary Policy Shocks under Rational Inattention." Discussion paper, Northwestern University.

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

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

[17] Van Nieuwerburgh, Stijn and Laura Veldkamp (2008): "Information Acquisition and Under-Diversification." Discussion paper, New York University.

[18] 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.

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

[20] Woodford, Michael (2008): "Information-Constrained State-Dependent Pricing." Discussion paper, Columbia University.

Figure 1: Impulse responses, benchmark economy

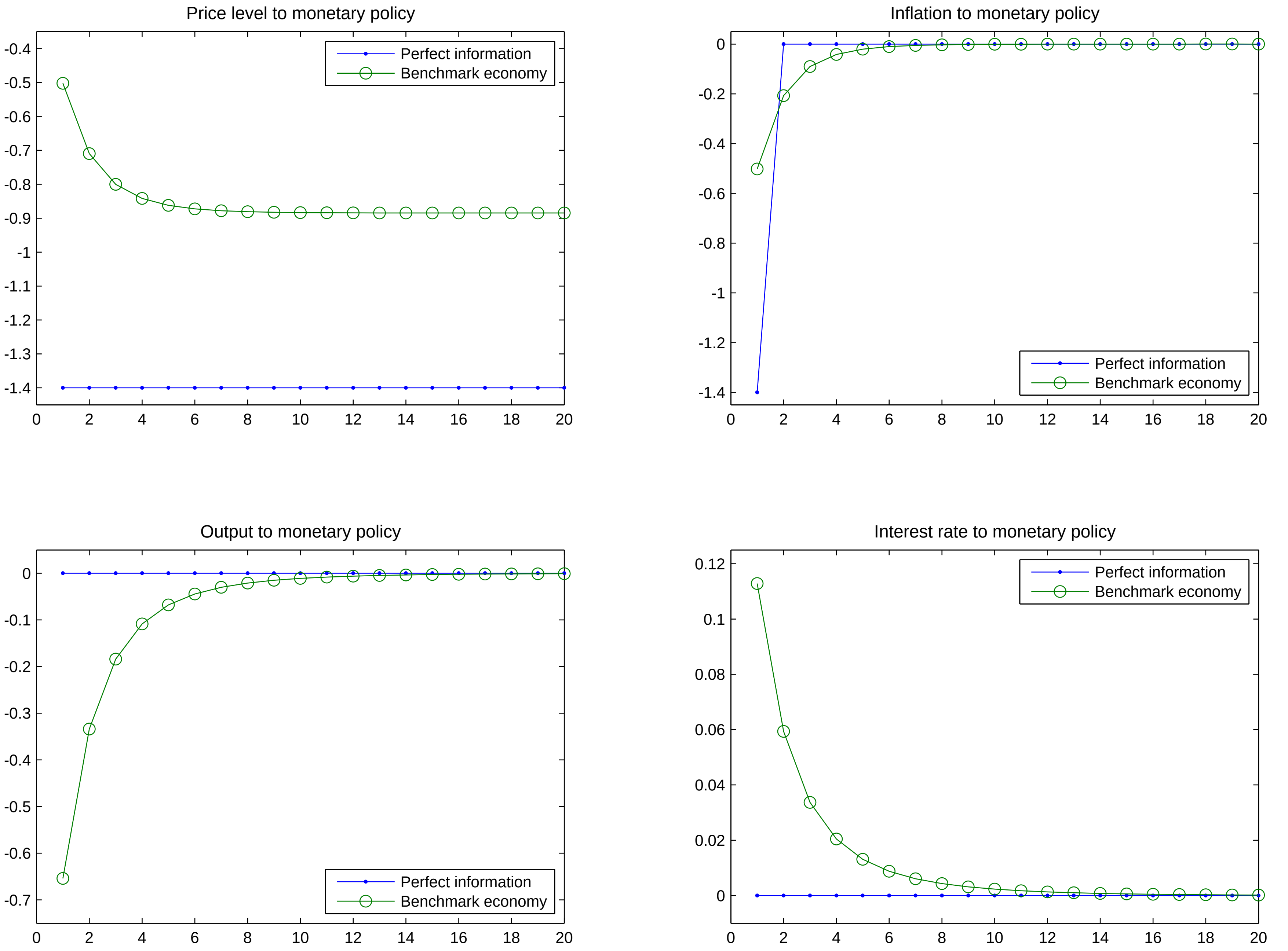

Figure 2: Impulse responses, benchmark economy

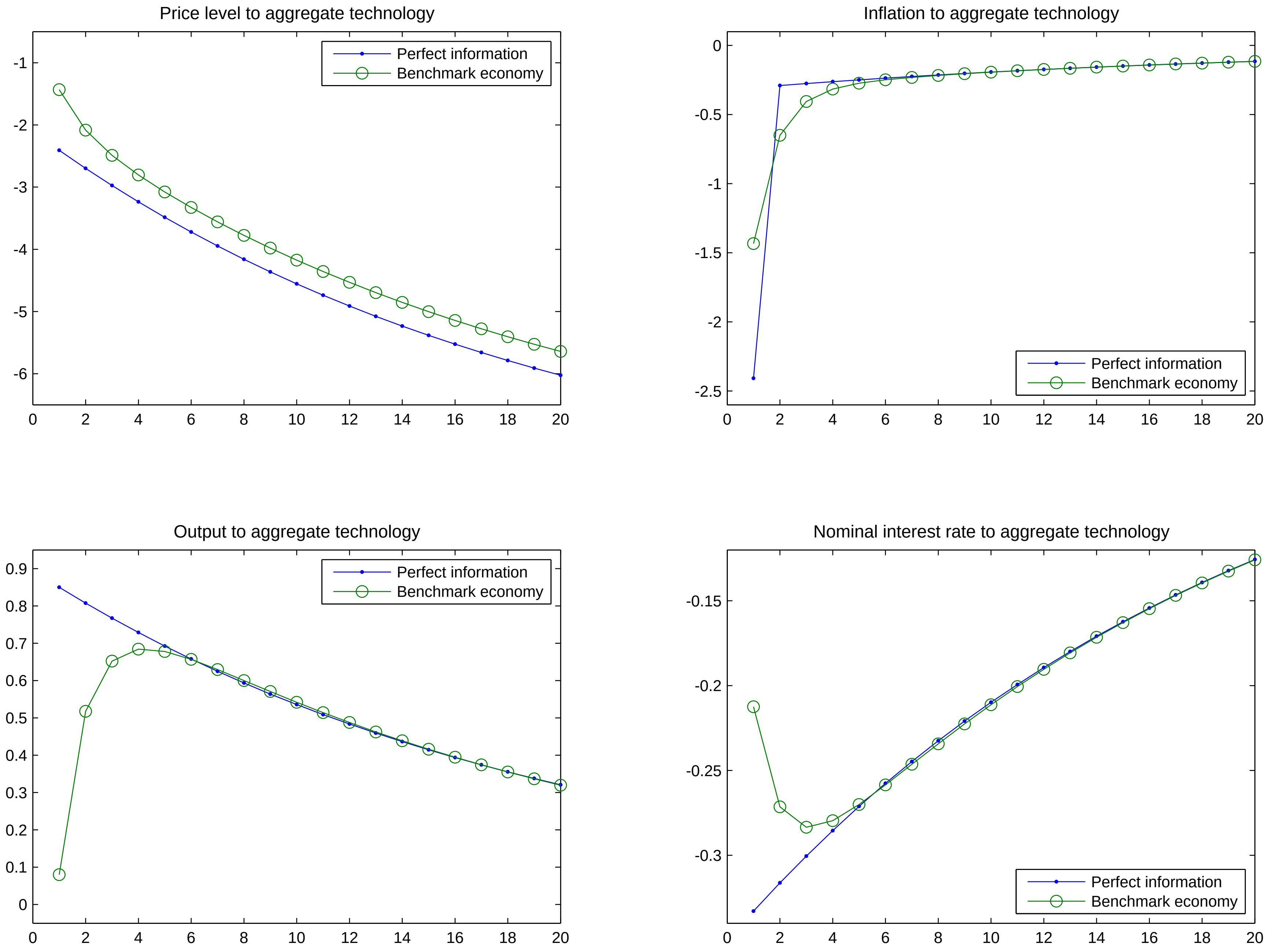

Figure 3: Impulse response of an individual price to a firm-specific productivity shock

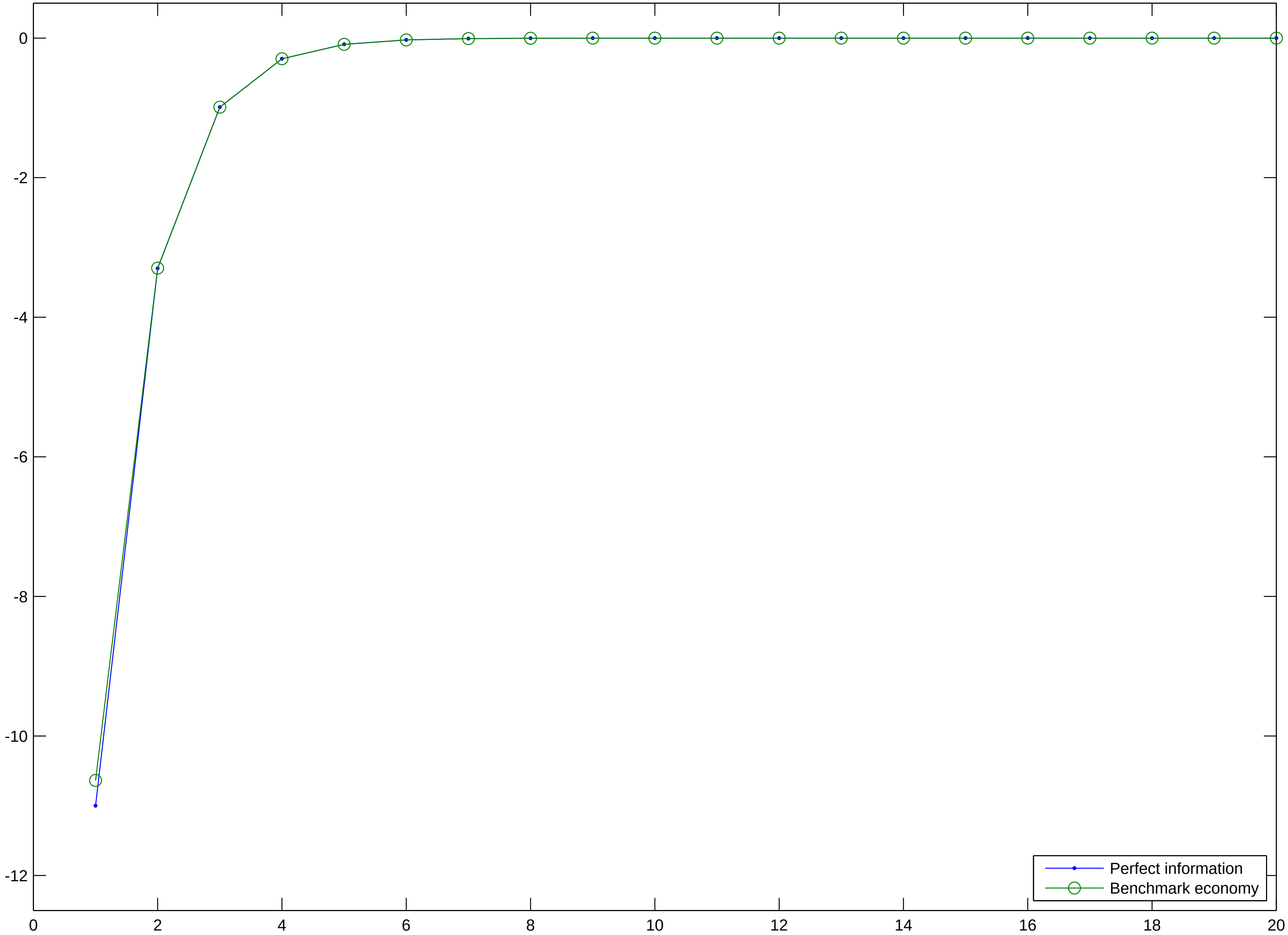

Figure 4: Impulse responses, benchmark economy and the Calvo model

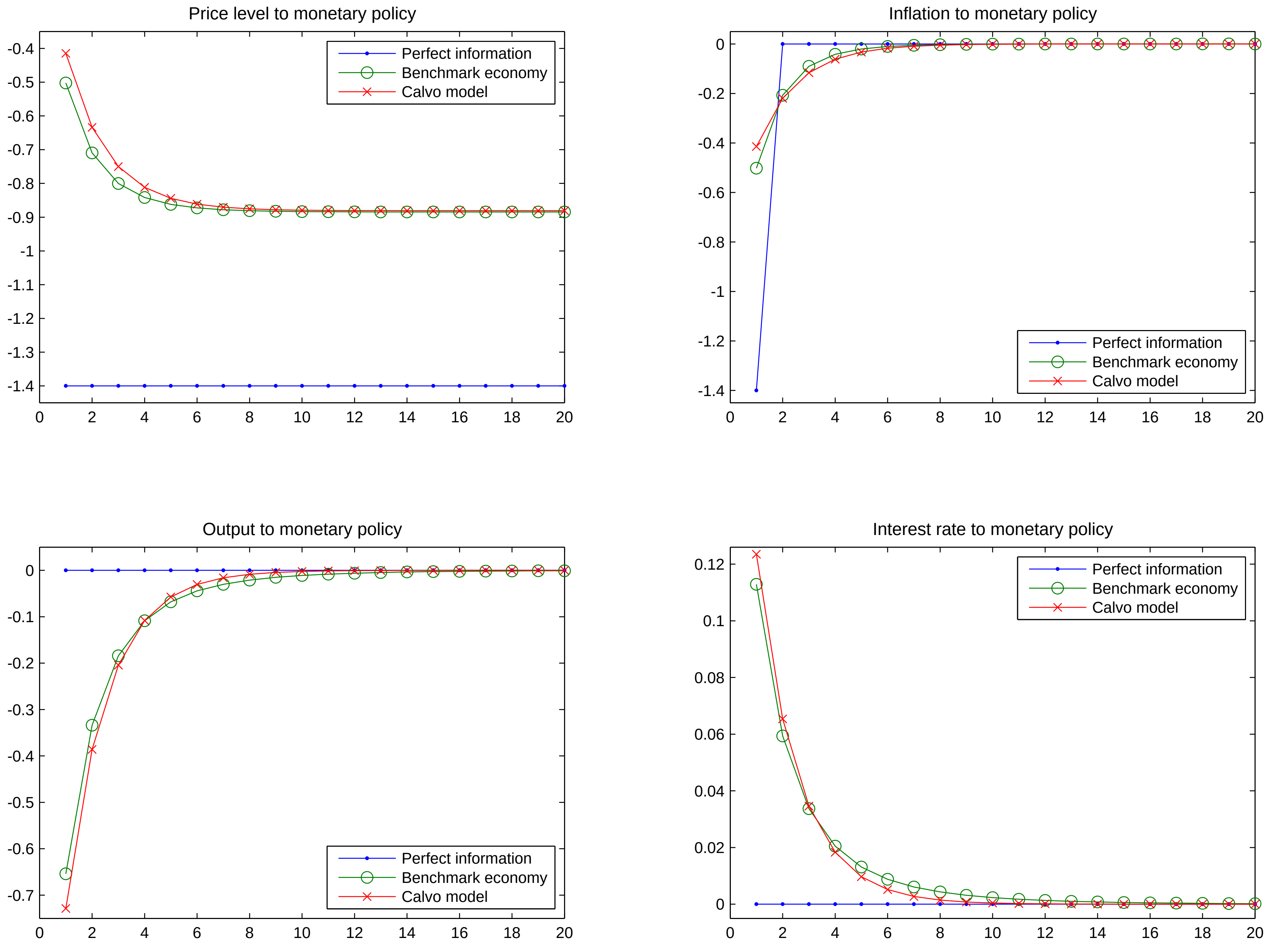

Figure 5: Impulse responses, benchmark economy and the Calvo model

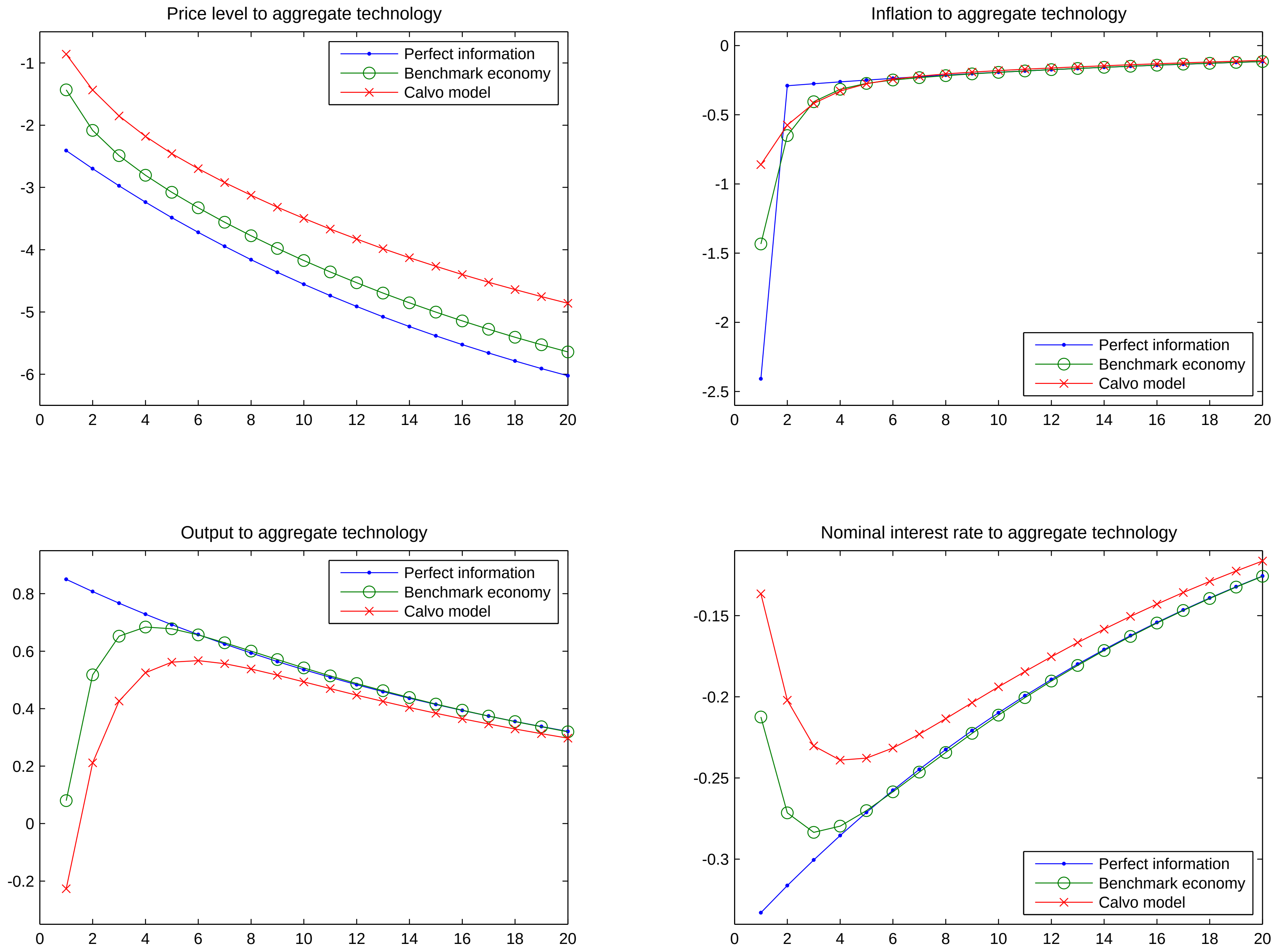

Figure 6: Impulse responses, benchmark economy and an economy with more real rigidity

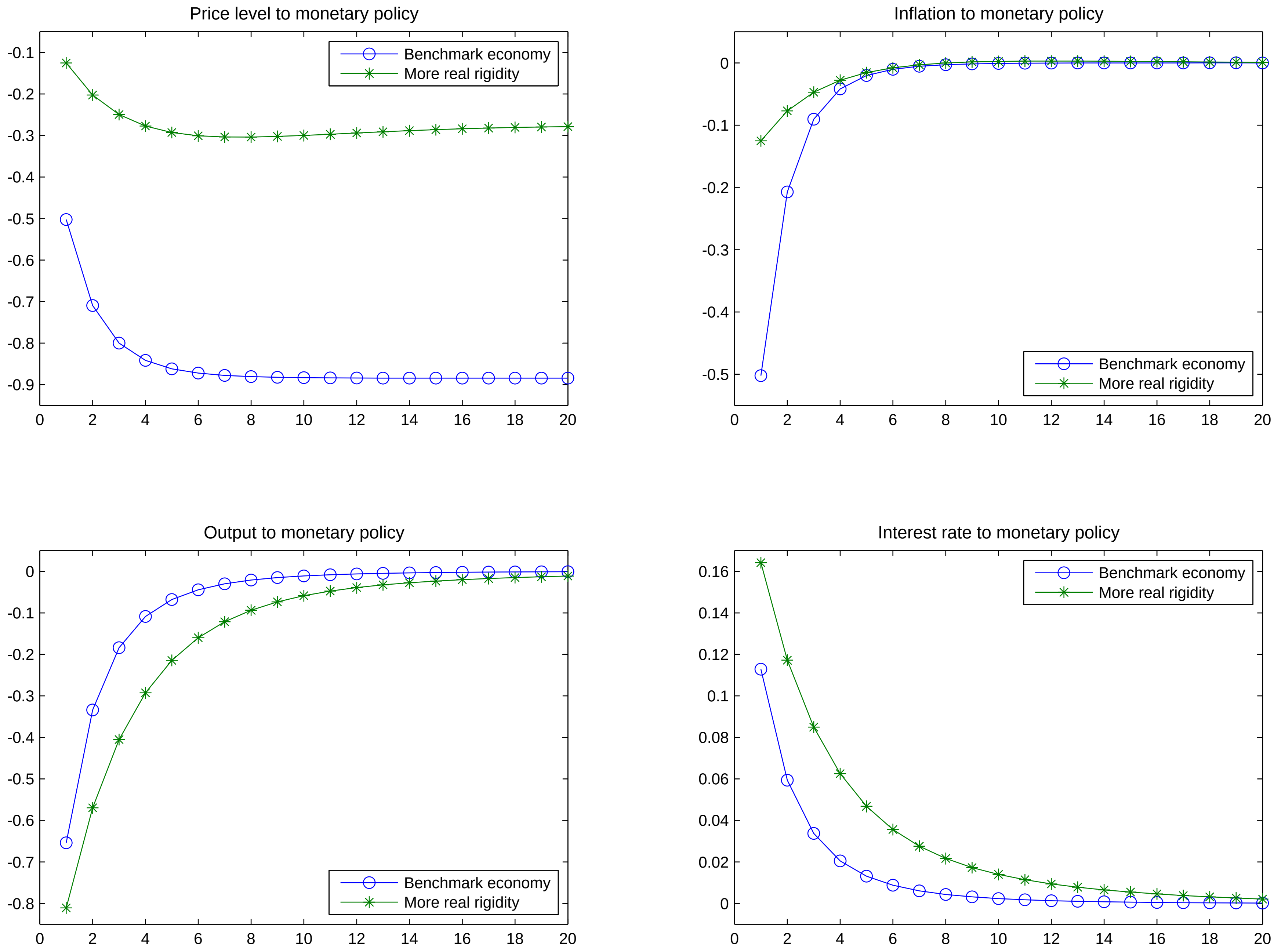

Figure 7: Impulse responses, benchmark economy and an economy with larger monetary policy shocks

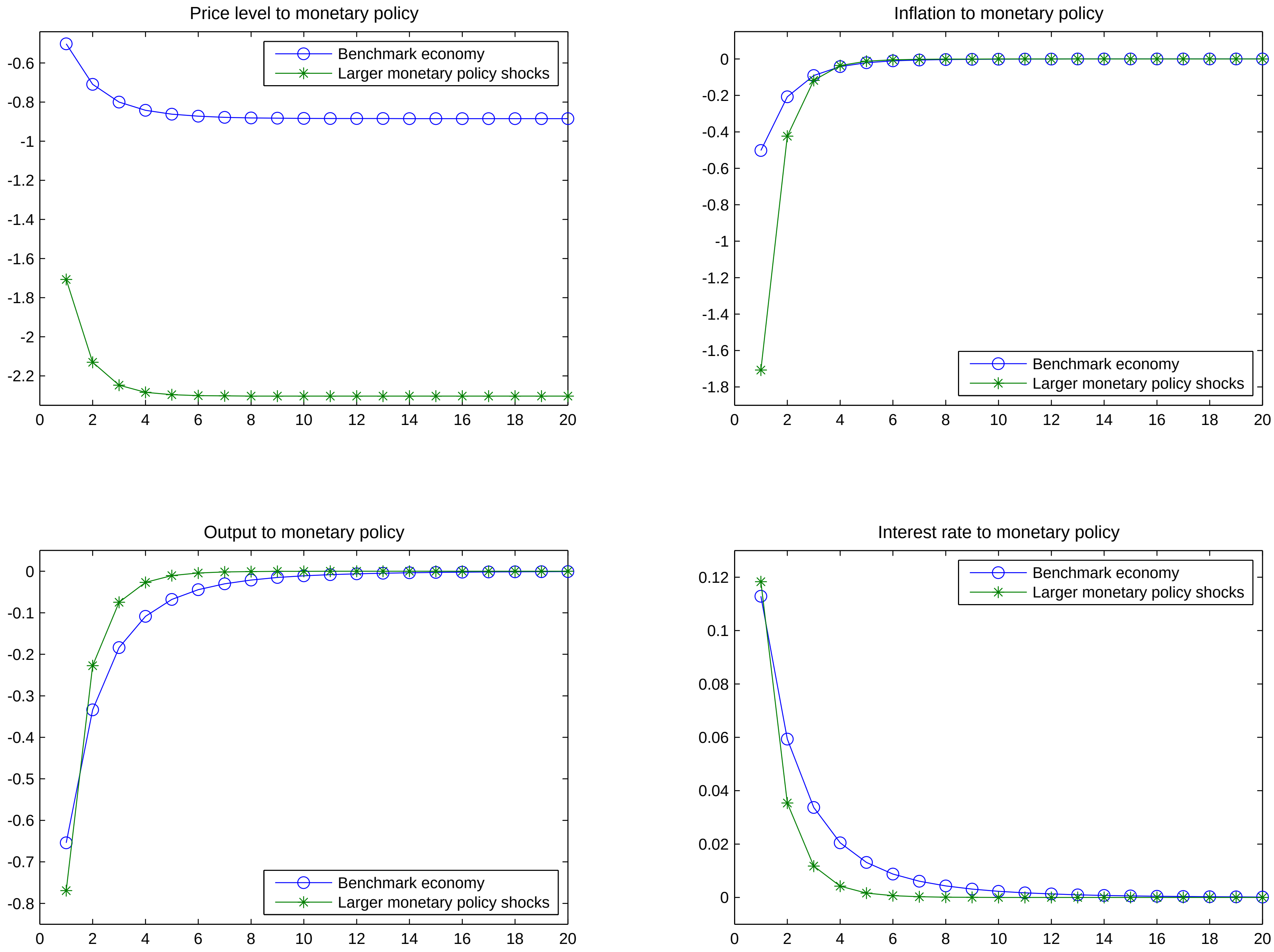

Figure 8: Impulse responses, household problem