Business Cycle Dynamics under Rational Inattention*

Bartosz Maćkowiak
European Central Bank and CEPR

Mirko Wiederholt
Northwestern University

First draft: June 2007. This draft: September 2009

Preliminary Version

Abstract

This paper develops a dynamic stochastic general equilibrium model with rational inattention. Households and decision-makers in firms have limited attention and decide how to allocate their attention. We study the implications of rational inattention for business cycle dynamics. We find that the impulse responses of prices under rational inattention have several properties of empirical impulse responses: (i) prices respond slowly to monetary policy shocks, (ii) prices respond faster to aggregate TFP shocks, and (iii) prices respond very fast to disaggregate shocks. As a result, 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 also find that consumption responds slowly to monetary policy shocks. For standard parameter values, deviations from the consumption Euler equation are cheap in utility terms, implying that households devote little attention to the consumption-saving decision.

*We thank for helpful comments: Paco Buera, Larry Christiano, James Costain, Martin Eichenbaum, Christian Hellwig, Giorgio Primiceri, Bruno Strulovici as well as seminar and conference participants at Bank of Canada, Chicago Fed, Columbia, Cowles Foundation Summer Conference 2009, DePaul, Duke, European Central Bank, ESSIM 2008, EUI, Harvard, MIT, Minneapolis Fed, Minnesota Workshop in Macroeconomic Theory 2009, NBER Summer Institute 2008, NAWMES 2008, Princeton, Richmond Fed, Riksbank, SED 2008, Stony Brook, Toronto, UCSD, University of Chicago and University of Hong Kong. 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 decision-makers have limited attention and decide how to allocate their attention. Following Sims (2003), we model attention as an information flow and we model limited attention as a constraint on information flow. We let agents choose the allocation of information flow. We study the implications of rational (in)attention for business cycle dynamics.

The economy consists of households, firms and a government. Firms produce differentiated goods using a variety of types of labor. Households supply the different types of labor, consume a variety of goods, and hold nominal government bonds. Decision-makers in firms take price setting and factor mix decisions. Households take consumption and wage setting decisions. The central bank sets the nominal interest rate according to a Taylor rule. Prices and wages are physically fully flexible and there is no habit formation in consumption. The only source of inertia in the model is the limited attention by decision-makers. We compute the impulse responses of prices and quantities to monetary policy shocks, aggregate technology shocks, and micro-level shocks under rational inattention by all decision-makers. We find that the model can match several features of empirical impulse responses.

We find that, in our model and for our parameter values, rational inattention by decision-makers in firms has the following implications. The price level responds slowly to monetary policy shocks. More precisely, the impulse response of the price level to monetary policy shocks under rational inattention by decision-makers in firms resembles the impulse response in a Calvo model with an average price duration of 7.5 months. At the same time, the price level responds fairly quickly to aggregate technology shocks, and individual prices respond very quickly to micro-level shocks. The reason is the optimal allocation of attention. Decision-makers in firms decide to devote little attention to monetary policy disturbances, about twice as much attention to the state of aggregate technology, and a lot of attention to market-specific conditions. Therefore, prices respond slowly to monetary policy shocks, prices respond fairly quickly to aggregate technology shocks, and prices respond very quickly to market-specific shocks.

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 of monetary policy shocks. More precisely, profit losses due to sub-optimal price responses to aggregate conditions are 23 times smaller than in the Calvo model; and profit losses due to sub-optimal price responses to idiosyncratic conditions are 57 times smaller than in the Calvo model that generates the same real effects of monetary policy shocks. The main reason for this result is the optimal allocation of attention, implying that prices respond slowly to monetary policy shocks, but prices respond fairly quickly to aggregate technology shocks, and prices respond very quickly to idiosyncratic shocks. By contrast, in the Calvo model prices respond slowly to all those shocks. The other reason for this result is that under rational inattention by decision-makers in firms deviations of the actual price from the profit-maximizing price are less likely to be extreme than in the Calvo model.

When we add rational inattention by households, we find that households devote little attention to the consumption-saving decision because for standard parameter values deviations from the consumption Euler equation are cheap in utility terms. Since households devote little attention to the consumption-saving decision, consumption responds slowly to shocks. It turns out that the impulse responses of consumption to shocks look similar to the impulse responses of consumption in a model with habit formation.

This paper is related to two strands of literature: (i) the literature on rational inattention (e.g. Sims (2003, 2006), Luo (2008), Maćkowiak and Wiederholt (2009), Van Nieuwerburgh and Veldkamp (2008), and Woodford (2009)); and (ii) the literature on business cycle models with imperfect information (e.g. Lucas (1972), Woodford (2002), Mankiw and Reis (2002), Lorenzoni (2008) and Angeletos and La'O (2009)). The main innovation with respect to the existing literature on rational inattention is that we solve a dynamic stochastic general equilibrium model. The main innovation with respect to the existing literature on business cycle models with imperfect information is that the information structure is the outcome of an optimization problem.

The paper is organized as follows. Section 2 describes all features of the economy apart from the information structure. Section 3 characterizes the steady state of the non-stochastic version of the economy. In Section 4 we derive the objective that decision-makers in firms maximize when they decide how to allocate their attention. In Section 5 we derive the objective that households maximize when they decide what to focus on. Section 6 describes issues related to aggregation. Section 7 characterizes the solution of the model under perfect information. Section 8 presents numerical solutions of the model under rational inattention by decision-makers in firms. Here we maintain the assumption that households have perfect information. Section 9 presents numerical solutions of the model under rational inattention by decision-makers in firms and households. Section 10 concludes.

2 Model

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

2.1 Households

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

Time is discrete and households have an infinite horizon. Each household seeks to maximize the expected discounted sum of period utility. The discount factor is $\beta \in (0,1)$. The period utility function is

$$U\left(C_{jt}, L_{jt}\right) = \frac{C_{jt}^{1-\gamma} - 1}{1-\gamma} - \varphi \frac{L_{jt}^{1+\psi}}{1+\psi}, \tag{1}$$

where

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

Here C_{ijt} is consumption of good i by household j in period t, C_{jt} is composite consumption by household j in period t and L_{jt} is labor supply by household j in period t. The parameter $\gamma > 0$ is the inverse of the intertemporal elasticity of substitution and the parameters $\varphi > 0$ and $\psi \geq 0$ affect the disutility of supplying labor. There are I different consumption goods and the parameter $\theta > 1$ is the elasticity of substitution between those consumption goods.[1]

[1] The assumption of a constant elasticity of substitution between consumption goods is only for ease of exposition. One could use a general constant returns-to-scale aggregator.

The flow budget constraint of household j in period t reads

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

Here P_{it} is the price of good i in period t, B_{jt} are holdings of nominal government bonds by household j between period t and period $t+1$, R_t is the nominal gross interest rate on those bond holdings, W_{jt} is the nominal wage rate for labor supplied by household j in period t, τ_w is a wage subsidy paid by the government, (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 households have the same initial bond holdings $B_{j,-1} > 0$. We also assume that bond holdings have to be positive in every period, $B_{jt} > 0$. We have to make some assumption to rule out Ponzi schemes. We choose this particular assumption because it allows us to rewrite the model in terms of logs of all variables. One can think of households having an account. The account holds only nominal government bonds, and the balance on the account has to be positive.

In every period, each household chooses a consumption vector, $(C_{1jt}, \ldots, C_{Ijt})$, and a wage rate, W_{jt}. Each household commits to supply any quantity of labor at that wage rate.

Each household takes as given: all prices of consumption goods, all wage rates set by other households, the nominal interest rate and all aggregate quantities.

2.2 Firms

There are I firms in the economy. Firms supply the differentiated consumption goods.

Firm i supplies good i. The production function of firm i is

$$Y_{it} = e^{a_t} e^{a_{it}} L_{it}^{\alpha}, \quad (4)$$

where

$$L_{it} = \left(\sum_{j=1}^{J} L_{ijt}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}. \quad (5)$$

Here Y_{it} is output, L_{ijt} is input of type j labor, L_{it} is composite labor input and $(e^{a_t} e^{a_{it}})$ is total factor productivity of firm i in period t. Total factor productivity has an aggregate component, e^{a_t}, and a firm-specific component, $e^{a_{it}}$. Type j labor is the labor supplied by household j. Since there are J households, there are J types of labor.[2] The parameter $\eta > 1$ is the elasticity of substitution between those types of labor. The parameter $\alpha \in (0,1]$ is the elasticity of output with respect to composite labor.

Nominal profits of firm i in period t equal

$$(1+\tau_p) P_{it} Y_{it} - \sum_{j=1}^{J} W_{jt} L_{ijt}, \tag{6}$$

where τ_p is a production subsidy paid by the government.

In every period, each firm sets a price, P_{it}, and chooses a factor mix, $\left(\hat{L}_{i1t}, \ldots, \hat{L}_{i(J-1)t}\right)$, where $\hat{L}_{ijt} = (L_{ijt}/L_{it})$ denotes firm i's relative input of type j labor in period t. Each firm commits to supply any quantity of the good at that price. Each firm then produces the quantity demanded with the chosen factor mix.

Each firm takes as given: all prices set by other firms, all wage rates set by households, the nominal interest rate and all aggregate quantities.

2.3 Government

There is a monetary authority and a fiscal authority. The monetary authority sets the nominal interest rate according to the rule

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

where $\Pi_t = (P_t/P_{t-1})$ is inflation, Y_t is aggregate output defined as

$$Y_t = \left(\sum_{i=1}^{I} P_{it} Y_{it}\right) / P_t, \tag{8}$$

and ε_t^R is a monetary policy shock. The price index P_t will be defined later. Here R, Π and Y denote 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

$$T_t + (B_t - B_{t-1}) = (R_{t-1} - 1) B_{t-1} + \tau_p \left(\sum_{i=1}^{I} P_{it} Y_{it}\right) + \tau_w \left(\sum_{j=1}^{J} W_{jt} L_{jt}\right). \tag{9}$$

[2] The assumption that all types of labor appear in the labor aggregator is for ease of exposition. One could assume that a firm-specific subset of types of labor appear in the labor aggregator.

The government has to finance interest on nominal government bonds, the production subsidy and the wage 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 arising from firms' market power in the goods market and households' market power in the labor market. In particular, we assume that

$$\tau_p = \frac{\vartheta}{\vartheta - 1} - 1, \tag{10}$$

where ϑ denotes the price elasticity of demand, and

$$\tau_w = \frac{\zeta}{\zeta - 1} - 1, \tag{11}$$

where ζ denotes the wage elasticity of labor demand.[3] We make this assumption to abstract from the level distortions arising from monopolistic competition.

2.4 Shocks

There are three types of shocks in the economy: aggregate technology shocks, firm-specific productivity shocks and monetary policy shocks. We assume that, for all $i = 1, \ldots, I$, the stochastic processes $\{a_t\}$, $\{a_{it}\}$ and $\left\{\varepsilon_t^R\right\}$ are independent. Furthermore, we assume that the firm-specific productivity processes, $\{a_{it}\}$, are independent across firms. In addition, we assume that the number of firms is sufficiently large so that

$$\frac{1}{I}\sum_{i=1}^{I} a_{it} = 0. \tag{12}$$

Finally, we assume that a_t follows a stationary Gaussian first-order autoregressive process with mean zero, each a_{it} follows a stationary Gaussian first-order autoregressive process with mean zero, and ε_t^R follows a Gaussian white noise process. In the following, we denote the period t innovation to a_t and a_{it} by ε_t^A and ε_{it}^I, respectively.

[3] When households have perfect information then $\vartheta = \theta$ and thus $\tau_p = \frac{\theta}{\theta-1} - 1$. By contrast, when households have imperfect information then ϑ (the price elasticity of demand) may differ from the parameter θ. Therefore, the value of the production subsidy (10) may vary across information structures. For the same reason, the value of the wage subsidy (11) may vary across information structures.

2.5 Notation

In this subsection, we introduce notation that will be convenient. Throughout the paper, C_t will denote aggregate composite consumption

$$C_t = \sum_{j=1}^{J} C_{jt}, \tag{13}$$

and L_t will denote aggregate composite labor input

$$L_t = \sum_{i=1}^{I} L_{it}. \tag{14}$$

Furthermore, $\hat{P}_{it}$ will denote the relative price of good i

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

and $\hat{W}_{jt}$ will denote the relative wage rate for type j labor

$$\hat{W}_{jt} = \frac{W_{jt}}{W_t}. \tag{16}$$

Finally, $\tilde{W}_{jt}$ will denote the real wage rate for type j labor

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

and $\tilde{W}_t$ will denote the real wage index

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

In each section, we will specify the definition of P_t and W_t.

3 Non-stochastic steady state

We begin by characterizing the non-stochastic steady state of the economy described in the previous section. We define a non-stochastic steady state as a solution of the non-stochastic version of the economy with the property that real quantities, 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, P_t denotes the following price index

$$P_t = \left(\sum_{i=1}^{I} P_{it}^{1-\theta} \right)^{\frac{1}{1-\theta}}, \tag{19}$$

and W_t denotes the following wage index

$$W_t = \left(\sum_{j=1}^{J} W_{jt}^{1-\eta} \right)^{\frac{1}{1-\eta}}. \tag{20}$$

In the non-stochastic steady state, the households' first-order conditions read

$$\frac{R}{\Pi} = \frac{1}{\beta}, \tag{21}$$

$$\frac{C_{ij}}{C_j} = \hat{P}_i^{-\theta}, \tag{22}$$

and

$$\tilde{W}_j = \varphi \left(\hat{W}_j^{-\eta} L \right)^{\psi} C_j^{\gamma}. \tag{23}$$

The firms' first-order conditions read

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

and

$$\hat{L}_{ij} = \hat{W}_j^{-\eta}. \tag{25}$$

The firms' price setting equation (24) implies that all firms set the same price in the non-stochastic steady state. Households therefore consume the different consumption goods in equal amounts, implying that all firms produce the same amount. Since in addition all firms have the same technology in the non-stochastic steady state, all firms have the same composite labor input. It follows from the definition of the price index (19), the consumption aggregator (2) and the definition of aggregate composite labor input (14) that

$$\hat{P}_i^{1-\theta} = \left(\frac{C_{ij}}{C_j} \right)^{\frac{\theta-1}{\theta}} = \frac{L_i}{L} = \frac{1}{I}. \tag{26}$$

Furthermore, in the non-stochastic version of the economy, all households face the same decision problem, have the same information and their decision problem has a unique constant solution, implying that all households choose the same consumption level and set the same wage rate in the non-stochastic steady state. Firms therefore hire the different types of labor in equal amounts. It follows from the definition of aggregate composite consumption (13), the definition of the wage index (20) and the labor aggregator (5) that

$$\frac{C_j}{C} = \hat{W}_j^{1-\eta} = \hat{L}_{ij}^{\frac{\eta-1}{\eta}} = \frac{1}{J}. \tag{27}$$

We will use equations (21)-(27) below.

One can show that equations (21)-(27), $Y_i = L_i^{\alpha}$, $Y_i = C_i$ and $C_i = \hat{P}_i^{-\theta} C$ imply that all variables appearing in equations (21)-(27) are uniquely determined apart from the nominal interest rate, R, and inflation, Π. For ease of exposition, we select $\Pi = 1$. Equation (21) then implies $R = (1/\beta)$. It is worth pointing out in this context that the steady-state inflation rate has no effect on real variables in both the non-stochastic and the stochastic version of the economy. In addition, in the non-stochastic steady state, the initial price level, P_{-1}, is not determined. We will assume that P_{-1} equals some value $\bar{P}_{-1}$. Finally, for given initial real bond holdings $\left(B_{j,-1}/\bar{P}_{-1}\right)$, fiscal variables in the non-stochastic steady state are uniquely determined by the requirement that real quantities are constant over time. The reason is that real bond holdings are a real quantity and real bond holdings are constant over time if and only if the government runs a balanced budget in real terms (i.e. real lump-sum taxes equal the sum of real interest payments and real subsidy payments).

4 Derivation of the firms' objective

In this section, we derive a log-quadratic approximation to the expected discounted sum of profits. We will use this expression below when we assume that decision-makers in firms choose the allocation of attention so as to maximize the expected discounted sum of profits. To derive this expression, we proceed in four steps: (i) we make a guess concerning the demand function for a consumption good, (ii) we substitute the demand function and the production function into the expression for profits to obtain the profit function, (iii) we make an assumption about how decision-makers in firms value profits in different states of the world, and (iv) we compute a log-quadratic approximation to the expected discounted sum of profits.

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

$$C_{it} = \varsigma \left(\frac{P_{it}}{P_t} \right)^{-\vartheta} C_t, \tag{28}$$

where C_t is aggregate composite consumption, P_t is a price index satisfying the following equation for some function d that is homogenous of degree one, symmetric and continuously differentiable

$$P_t = d\left(P_{1t}, \ldots, P_{It}\right), \tag{29}$$

and $\vartheta > 1$ and $\varsigma > 0$ are undetermined coefficients satisfying

$$\varsigma \hat{P}_i^{-\vartheta} = \hat{P}_i^{-\theta}. \tag{30}$$

When we solve the model for alternative assumptions about the information structure below, we will always verify that this guess concerning the demand function is correct.[4]

Second, we substitute the production function (4)-(5) and the demand function (28) into the expression for nominal profits (6) to obtain the profit function. We begin by rewriting the expression for nominal profits (6)

$$(1+\tau_p) P_{it} Y_{it} - \sum_{j=1}^{J} W_{jt} L_{ijt} = (1+\tau_p) P_{it} Y_{it} - L_{it} \left[\sum_{j=1}^{J} W_{jt} \hat{L}_{ijt} \right], \tag{31}$$

where $\hat{L}_{ijt} = (L_{ijt}/L_{it})$ is firm i's relative input of type j labor. The term in square brackets on the right-hand side is the wage bill per unit of composite labor input. Rearranging equations (4)-(5) yields

$$L_{it} = \left(\frac{Y_{it}}{e^{a_t} e^{a_{it}}} \right)^{\frac{1}{\alpha}}, \tag{32}$$

and

$$1 = \sum_{j=1}^{J} \hat{L}_{ijt}^{\frac{\eta-1}{\eta}}. \tag{33}$$

Substituting the technology (32)-(33), $Y_{it} = C_{it}$ and the demand function (28) into the expression for nominal profits (31) yields the profit function

$$(1+\tau_p) P_{it} \varsigma \left(\frac{P_{it}}{P_t} \right)^{-\vartheta} C_t - \left[\frac{\varsigma \left(\frac{P_{it}}{P_t} \right)^{-\vartheta} C_t}{e^{a_t} e^{a_{it}}} \right]^{\frac{1}{\alpha}} \left[\sum_{j=1}^{J-1} W_{jt} \hat{L}_{ijt} + W_{Jt} \left(1 - \sum_{j=1}^{J-1} \hat{L}_{ijt}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \right]. \tag{34}$$

[4] To give the simplest example, when households have perfect information then P_t is given by equation (19), $\vartheta = \theta$ and $\varsigma = 1$.

Profits of firm i in period t depend on the following variables that the decision-maker in the firm chooses: $P_{it}, \hat{L}_{i1t}, \ldots, \hat{L}_{i(J-1)t}$; and on the following variables that the decision-maker in the firm takes as given: $P_t, a_t, a_{it}, C_t, W_{1t}, \ldots, W_{Jt}$.

Third, we make an assumption about how decision-makers in firms value profits in different states of the world. Since the economy described in Section 2 is an incomplete-markets economy with multiple owners of a firm, it is unclear how firms should value profits in different states of the world. Therefore, we assume a general stochastic discount factor. In particular, we assume that, in period -1, decision-makers in firms value nominal profits in period t using the following stochastic discount factor

$$Q_{-1,t} = \beta^t \lambda \left(C_{1t}, \ldots, C_{Jt}\right) \frac{1}{P_t}, \tag{35}$$

where P_t is the price index appearing in the demand function (28) and λ is some twice continuously differentiable function with the property[5]

$$\lambda \left(C_1, \ldots, C_J\right) = C_j^{-\gamma}. \tag{36}$$

Then, in period -1, the expected discounted sum of profits equals

$$E_{i,-1} \left[\sum_{t=0}^{\infty} \beta^t F \left(\hat{P}_{it}, \hat{L}_{i1t}, \ldots, \hat{L}_{i(J-1)t}, a_t, a_{it}, C_{1t}, \ldots, C_{Jt}, \tilde{W}_{1t}, \ldots, \tilde{W}_{Jt} \right) \right], \tag{37}$$

where $E_{i,-1}$ is the expectation operator conditioned on the information of the decision-maker of firm i in period -1 and the function F is given by

$$\begin{aligned} & F \left(\hat{P}_{it}, \hat{L}_{i1t}, \ldots, \hat{L}_{i(J-1)t}, a_t, a_{it}, C_{1t}, \ldots, C_{Jt}, \tilde{W}_{1t}, \ldots, \tilde{W}_{Jt} \right) \\ = \; & \lambda \left(C_{1t}, \ldots, C_{Jt}\right) \left(1+\tau_p\right) \varsigma \hat{P}_{it}^{1-\vartheta} \left(\sum_{j=1}^{J} C_{jt} \right) \\ & -\lambda \left(C_{1t}, \ldots, C_{Jt}\right) \left[\frac{\varsigma \hat{P}_{it}^{-\vartheta} \left(\sum_{j=1}^{J} C_{jt} \right)}{e^{a_t} e^{a_{it}}} \right]^{\frac{1}{\alpha}} \left[\sum_{j=1}^{J-1} \tilde{W}_{jt} \hat{L}_{ijt} + \tilde{W}_{Jt} \left(1 - \sum_{j=1}^{J-1} \hat{L}_{ijt}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \right] \end{aligned} \tag{38}$$

[5] For example, the stochastic discount factor could be a weighted average of the marginal utilities of the different households (i.e. $\lambda \left(C_{1t}, \ldots, C_{Jt}\right) = \sum_{j=1}^{J} \lambda_j C_{jt}^{-\gamma}$ with $\lambda_j \geq 0$ and $\sum_{j=1}^{J} \lambda_j = 1$). Equation (36) would be satisfied because all households have the same marginal utility in the non-stochastic steady state.

In the following, small variables denote log-deviations from the non-stochastic steady state. For example, $c_{jt} = \ln(C_{jt}/C_j)$. Expressing the real profit function F in terms of log-deviations from the non-stochastic steady state and using equations (10), (24)-(25), (27), (30), $Y_i = L_i^\alpha$, $Y_i = C_i$ and $C_i = \hat{P}_i^{-\theta} C$ yields the following expression for the expected discounted sum of profits

$$E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t f\left(\hat{p}_{it}, \hat{l}_{i1t}, \ldots, \hat{l}_{i(J-1)t}, a_t, a_{it}, c_{1t}, \ldots, c_{Jt}, \tilde{w}_{1t}, \ldots, \tilde{w}_{Jt}\right)\right], \tag{39}$$

where

$$\begin{aligned}
& f\left(\hat{p}_{it}, \hat{l}_{i1t}, \ldots, \hat{l}_{i(J-1)t}, a_t, a_{it}, c_{1t}, \ldots, c_{Jt}, \tilde{w}_{1t}, \ldots, \tilde{w}_{Jt}\right) \\
= \; & \lambda\left(C_1 e^{c_{1t}}, \ldots, C_J e^{c_{Jt}}\right)\frac{\vartheta}{\vartheta-1}\frac{1}{\alpha}\tilde{W}L_i\frac{1}{J}\sum_{j=1}^{J} e^{(1-\vartheta)\hat{p}_{it}+c_{jt}} \\
& -\lambda\left(C_1 e^{c_{1t}}, \ldots, C_J e^{c_{Jt}}\right)\tilde{W}L_i e^{-\frac{\vartheta}{\alpha}\hat{p}_{it}-\frac{1}{\alpha}(a_t+a_{it})}\left(\frac{1}{J}\sum_{j=1}^{J} e^{c_{jt}}\right)^{\frac{1}{\alpha}} \\
& \frac{1}{J}\left[\sum_{j=1}^{J-1} e^{\tilde{w}_{jt}+\hat{l}_{ijt}} + e^{\tilde{w}_{Jt}}\left(J - \sum_{j=1}^{J-1} e^{\frac{\eta-1}{\eta}\hat{l}_{ijt}}\right)^{\frac{\eta}{\eta-1}}\right].
\end{aligned} \tag{40}$$

Fourth, we compute a log-quadratic approximation to the expected discounted sum of profits around the non-stochastic steady state. We obtain the following result.

Proposition 1 *(Expected discounted sum of profits) Let f denote the real profit function defined by equation (40) and let $\tilde{f}$ denote the second-order Taylor approximation to f at the non-stochastic steady state. Let $E_{i,-1}$ denote the expectation operator conditioned on the information of the decision-maker of firm i in period -1. Let x_t, z_t and v_t denote the following vectors*

$$x_t = \begin{pmatrix} \hat{p}_{it} & \hat{l}_{i1t} & \cdots & \hat{l}_{i(J-1)t} \end{pmatrix}', \tag{41}$$

$$z_t = \begin{pmatrix} a_t & a_{it} & c_{1t} & \cdots & c_{Jt} & \tilde{w}_{1t} & \cdots & \tilde{w}_{Jt} \end{pmatrix}', \tag{42}$$

$$v_t = \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}'. \tag{43}$$

Let $v_{m,t}$ denote the mth element of v_t. Suppose that there exist two constants $\delta < (1/\beta)$ and $A \in \mathbb{R}$ such that, for each period $t \geq 0$ and for all m and n,

$$E_{i,-1}\left|v_{m,t}v_{n,t}\right| < \delta^t A. \tag{44}$$

Then

$$E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^{t}\tilde{f}\left(x_{t},z_{t}\right)\right]-E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^{t}\tilde{f}\left(x_{t}^{*},z_{t}\right)\right]$$
$$=\sum_{t=0}^{\infty}\beta^{t}E_{i,-1}\left[\frac{1}{2}\left(x_{t}-x_{t}^{*}\right)'H\left(x_{t}-x_{t}^{*}\right)\right], \tag{45}$$

where the matrix H is given by

$$H=-C_{j}^{-\gamma}\tilde{W}L_{i}\begin{bmatrix}\frac{\vartheta}{\alpha}\left(1+\frac{1-\alpha}{\alpha}\vartheta\right) & 0 & \cdots & \cdots & 0\\ 0 & \frac{2}{\eta J} & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J}\\ \vdots & \frac{1}{\eta J} & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & \frac{1}{\eta J}\\ 0 & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J} & \frac{2}{\eta J}\end{bmatrix}, \tag{46}$$

and the vector x_t^* is given by:

$$\hat{p}_{it}^{*}=\frac{\frac{1-\alpha}{\alpha}}{1+\frac{1-\alpha}{\alpha}\vartheta}\left(\frac{1}{J}\sum_{j=1}^{J}c_{jt}\right)+\frac{1}{1+\frac{1-\alpha}{\alpha}\vartheta}\left(\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}\right)-\frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\vartheta}\left(a_{t}+a_{it}\right), \tag{47}$$

and

$$\hat{l}_{ijt}^{*}=-\eta\left(\tilde{w}_{jt}-\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}\right). \tag{48}$$

Proof. See Appendix A. ■

After the log-quadratic approximation to the real profit function, the profit-maximizing price in period t is given by equation (47) and the profit-maximizing factor mix in period t is given by equation (48). Furthermore, after the log-quadratic approximation to the real profit function, the loss in profits in period t in the case of a deviation from the profit-maximizing decisions (i.e. $x_t \neq x_t^*$) is given by the quadratic form in expression (45). The upper-left element of the matrix H determines the profit loss in the case of a sub-optimal price setting decision. The profit loss in the case of a sub-optimal price setting decision is increasing in the price elasticity of demand, ϑ, and increasing in the degree of decreasing returns-to-scale, $(1/\alpha)$. The lower-right block of the matrix H determines the profit loss in the case of a sub-optimal factor mix decision. The profit loss in the case of a sub-optimal factor mix decision is decreasing in the elasticity of substitution between types of labor, η, and depends on the number of types of labor, J. Note that the diagonal elements of H determine the profit loss in the case of a deviation in a single variable, while the off-diagonal elements of H determine how a deviation in one variable affects the loss in profits due to a deviation in another variable. Finally, condition (44) ensures that, in the expression for the expected discounted sum of profits, after the log-quadratic approximation to the real profit function, one can change the order of integration and summation and the infinite sum converges.

It is worth pointing out that the profit-maximizing decision vector (47)-(48) and the expected discounted sum of profit losses (45) depend only to a limited extent on the function λ appearing in the discount factor (35). The profit-maximizing decision vector (47)-(48) does not depend at all on the function λ because the profit-maximizing price and the profit-maximizing factor mix are the solution to a static decision problem in the economy described in Section 2. The expected discounted sum of profit losses (45) depends only on the steady-state value of the function λ because of the log-quadratic approximation to the real profit function.

So far we have only derived an expression for the expected discounted sum of profits in the economy described in Section 2, but from this expression one can already see to some extent how a decision-maker in a firm who cannot attend perfectly to all available information will allocate his/her attention. For example, the attention devoted to the price setting decision will depend on the profit loss that the firm incurs in the case of a price setting mistake (i.e., a deviation of the actual price from the profit-maximizing price). Thus, the attention devoted to the price setting decision will depend on the upper-left element of the matrix H. Furthermore, the decision-maker will track closely those changes in the environment that in expectation cause most of the variation in the profit-maximizing decisions. As one can see from equations (47)-(48), which changes in the environment in expectation cause most of the variation in the profit-maximizing decisions depends on the technology parameters α and η, the calibration of the exogenous processes as well as the behavior of the other agents in the economy.

5 Derivation of the households' objective

In this section, we derive a log-quadratic approximation to the expected discounted sum of period utility. We will use this expression below when we assume that households choose the allocation of attention so as to maximize the expected discounted sum of period utility. To derive this expression, we proceed in three steps: (i) we make a guess concerning the demand function for a particular type of labor, (ii) we substitute the labor demand function, the consumption aggregator and the flow budget constraint into the period utility function to obtain a period utility function that incorporates those constraints, and (iii) we compute a log-quadratic approximation to the expected discounted sum of period utility.

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

$$L_{jt} = \xi \left(\frac{W_{jt}}{W_t} \right)^{-\zeta} L_t, \tag{49}$$

where L_t is aggregate composite labor input, W_t is a wage index satisfying the following equation for some function h that is homogenous of degree one, symmetric and continuously differentiable

$$W_t = h\left(W_{1t}, \ldots, W_{Jt}\right), \tag{50}$$

and $\zeta > 1$ and $\xi > 0$ are undetermined coefficients satisfying

$$\xi \hat{W}_j^{-\zeta} = \hat{W}_j^{-\eta}. \tag{51}$$

When we solve the model for alternative assumptions about the information structure below, we will always verify that this guess concerning the labor demand function is correct.[6]

Second, we substitute the consumption aggregator (2), the flow budget constraint (3) and the labor demand function (49) into the period utility function (1) to obtain a period utility function that incorporates those constraints. We begin by rewriting the flow budget constraint (3) as

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

[6] To give the simplest example, when firms have perfect information then W_t is given by equation (20), $\zeta = \eta$ and $\xi = 1$.

where $\hat{C}_{ijt} = (C_{ijt}/C_{jt})$ is relative consumption of good i by household j. The term in brackets on the left-hand side is consumption expenditure per unit of composite consumption. Rearranging yields

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

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

$$C_{jt} = \frac{\frac{R_{t-1}}{\Pi_t}\tilde{B}_{jt-1} - \tilde{B}_{jt} + (1+\tau_w) \tilde{W}_{jt}L_{jt} + \frac{\tilde{D}_t}{J} - \frac{\tilde{T}_t}{J}}{\sum_{i=1}^{I} \hat{P}_{it}\hat{C}_{ijt}}, \tag{52}$$

where $\tilde{B}_{jt} = (B_{jt}/P_t)$ are real bond holdings by the household, $\tilde{D}_t = (D_t/P_t)$ 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

$$1 = \sum_{i=1}^{I} \hat{C}_{ijt}^{\frac{\theta-1}{\theta}}. \tag{53}$$

Substituting the flow budget constraint (52), the consumption aggregator (53) and the labor demand function (49) into the period utility function (1) yields a period utility function that incorporates those constraints:

$$\frac{1}{1-\gamma}\left(\frac{\frac{R_{t-1}}{\Pi_t}\tilde{B}_{jt-1} - \tilde{B}_{jt} + (1+\tau_w)\tilde{W}_{jt}\xi\left(\frac{\tilde{W}_{jt}}{\tilde{W}_t}\right)^{-\zeta} L_t + \frac{\tilde{D}_t}{J} - \frac{\tilde{T}_t}{J}}{\sum_{i=1}^{I-1}\hat{P}_{it}\hat{C}_{ijt} + \hat{P}_{It}\left(1 - \sum_{i=1}^{I-1}\hat{C}_{ijt}^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}}\right)^{1-\gamma}$$
$$-\frac{1}{1-\gamma} - \frac{\varphi}{1+\psi}\left[\xi\left(\frac{\tilde{W}_{jt}}{\tilde{W}_t}\right)^{-\zeta} L_t\right]^{1+\psi}. \tag{54}$$

Expressing the period utility function (54) in terms of log-deviations from the non-stochastic steady state and using equations (11), (21)-(23), (26), (51) and $L_j = \hat{W}_j^{-\eta} L$ yields our final period utility function:

$$\frac{C_j^{1-\gamma}}{1-\gamma}\left(\frac{\frac{\omega_B}{\beta}e^{r_{t-1}-\pi_t+\tilde{b}_{jt-1}}-\omega_B e^{\tilde{b}_{jt}}+\frac{\zeta}{\zeta-1}\omega_W e^{(1-\zeta)\tilde{w}_{jt}+\zeta\tilde{w}_t+l_t}+\omega_D e^{\tilde{d}_t}-\omega_T e^{\tilde{t}_t}}{\frac{1}{I}\sum_{i=1}^{I-1}e^{\hat{p}_{it}+\hat{c}_{ijt}}+\frac{1}{I}e^{\hat{p}_{It}}\left(I-\sum_{i=1}^{I-1}e^{\frac{\theta-1}{\theta}\hat{c}_{ijt}}\right)^{\frac{\theta}{\theta-1}}}\right)^{1-\gamma}$$
$$-\frac{1}{1-\gamma}-\frac{C_j^{1-\gamma}}{1+\psi}\omega_W e^{-\zeta(1+\psi)(\tilde{w}_{jt}-\tilde{w}_t)+(1+\psi)l_t}, \tag{55}$$

where ω_B, ω_W, ω_D and ω_T denote the following steady-state ratios:

$$\begin{pmatrix}\omega_B & \omega_W & \omega_D & \omega_T\end{pmatrix}=\begin{pmatrix}\frac{\tilde{B}_j}{C_j} & \frac{\tilde{W}_j L_j}{C_j} & \frac{\frac{\tilde{D}}{J}}{C_j} & \frac{\frac{\tilde{T}}{J}}{C_j}\end{pmatrix}. \tag{56}$$

Third, we compute a log-quadratic approximation to the expected discounted sum of period utility around the non-stochastic steady state.

Proposition 2 *(Expected discounted sum of period utility) Let g denote the functional that is obtained by multiplying the period utility function (55) by β^t and summing over all t from zero to infinity. Let $\tilde{g}$ denote the second-order Taylor approximation to g at the non-stochastic steady state. Let $E_{j,-1}$ denote the expectation operator conditioned on information of household j in period -1. Let x_t, z_t and v_t denote the following vectors*

$$x_t = \begin{pmatrix}\tilde{b}_{jt} & \tilde{w}_{jt} & \hat{c}_{1jt} & \cdots & \hat{c}_{I-1jt}\end{pmatrix}', \tag{57}$$

$$z_t = \begin{pmatrix}r_{t-1} & \pi_t & \tilde{w}_t & l_t & \tilde{d}_t & \tilde{t}_t & \hat{p}_{1t} & \cdots & \hat{p}_{It}\end{pmatrix}', \tag{58}$$

$$v_t = \begin{pmatrix}x_t' & z_t' & 1\end{pmatrix}'. \tag{59}$$

Let $v_{m,t}$ denote the mth element of v_t. Suppose that

$$E_{j,-1}\left[\tilde{b}_{j,-1}^2\right]<\infty, \tag{60}$$

and, for all m,

$$E_{j,-1}\left|\tilde{b}_{j,-1}v_{m,0}\right|<\infty. \tag{61}$$

Furthermore, suppose that there exist two constants $\delta<(1/\beta)$ and $A\in\mathbb{R}$ such that, for each period $t\geq 0$, for $\tau=0,1$ and for all m and n,

$$E_{j,-1}\left|v_{m,t}v_{n,t+\tau}\right|<\delta^t A. \tag{62}$$

Then

$$
\begin{aligned}
& E_{j,-1}\left[\tilde{g}\left(\tilde{b}_{j,-1}, x_0, z_0, x_1, z_1, \ldots\right)\right] - E_{j,-1}\left[\tilde{g}\left(\tilde{b}_{j,-1}, x_0^*, z_0, x_1^*, z_1, \ldots\right)\right] \\
= & \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2}\left(x_t - x_t^*\right)' H_0 \left(x_t - x_t^*\right) + \left(x_t - x_t^*\right)' H_1 \left(x_{t+1} - x_{t+1}^*\right)\right]. \qquad (63)
\end{aligned}
$$

Here the matrix H_0 is given by

$$
H_0 = -C_j^{1-\gamma} \begin{bmatrix} \gamma\omega_B^2\left(1+\frac{1}{\beta}\right) & \gamma\omega_B\zeta\omega_W & 0 & \cdots & 0 \\ \gamma\omega_B\zeta\omega_W & \zeta\omega_W\left(\gamma\zeta\omega_W + 1 + \zeta\psi\right) & 0 & \cdots & 0 \\ 0 & 0 & \frac{2}{\theta I} & \cdots & \frac{1}{\theta I} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \frac{1}{\theta I} & \cdots & \frac{2}{\theta I} \end{bmatrix}, \qquad (64)
$$

and the matrix H_1 is given by

$$
H_1 = C_j^{1-\gamma} \begin{bmatrix} \gamma\omega_B^2 & \gamma\omega_B\zeta\omega_W & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}. \qquad (65)
$$

Furthermore, the process $\{x_t^*\}$ is defined by the following two requirements: (i) the vector v_t with $x_t = x_t^*$ satisfies conditions (60)-(62), and (ii) in each period $t \geq 0$,

$$
\begin{aligned}
c_{jt}^* &= E_t\left[-\frac{1}{\gamma}\left(r_t - \pi_{t+1} - \frac{1}{I}\sum_{i=1}^{I}\left(\hat{p}_{it+1} - \hat{p}_{it}\right)\right) + c_{jt+1}^*\right], \qquad (66) \\
\tilde{w}_{jt}^* &= \frac{\gamma}{1+\zeta\psi}c_{jt}^* + \frac{\psi}{1+\zeta\psi}\left(\zeta\tilde{w}_t + l_t\right) + \frac{1}{1+\zeta\psi}\left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \qquad (67) \\
\hat{c}_{ijt}^* &= -\theta\left(\hat{p}_{it} - \frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \qquad (68)
\end{aligned}
$$

where c_{jt}^* is defined by

$$
\begin{aligned}
c_{jt}^* = & \frac{\omega_B}{\beta}\left(r_{t-1} - \pi_t + \tilde{b}_{jt-1}^*\right) - \omega_B\tilde{b}_{jt}^* + \frac{\zeta}{\zeta-1}\omega_W\left[\left(1-\zeta\right)\tilde{w}_{jt}^* + \zeta\tilde{w}_t + l_t\right] \\
& + \omega_D\tilde{d}_t - \omega_T\tilde{t}_t - \left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \qquad (69)
\end{aligned}
$$

and E_t denotes the expectation operator conditioned on the entire history of the economy up to and including period t.

Proof. See Appendix B. ■

After the log-quadratic approximation to the expected discounted sum of period utility, stochastic processes for real bond holdings, the real wage rate and the consumption mix that satisfy conditions (60)-(62) can be ranked using equation (63). Equations (66)-(69) characterize the optimal behavior under perfect information (i.e. the decisions the household would take if in each period $t \geq 0$ the household knew the entire history of the economy up to and including period t). Equation (63) gives the loss in expected lifetime utility in the case of deviations from the optimal behavior under perfect information. The upper-left block of the matrix H_0 and the upper-left block of the matrix H_1 determine the loss in expected lifetime utility in the case of sub-optimal real bond holdings and wage setting. A single percentage deviation in real bond holdings from optimal bond holdings causes a larger utility loss the larger γ, ω_B and $(R/\Pi) = (1/\beta)$. See the (1,1) element of the matrix H_0. A single percentage deviation in the real wage rate from the optimal wage rate causes a larger utility loss the larger γ, ψ, ω_W and ζ. See the (2,2) element of the matrix H_0. Furthermore, the off-diagonal elements of H_0 show that a bond deviation in period t affects the utility cost of a wage deviation in period t, and the first row of H_1 shows that a bond deviation in period t affects both the utility cost of a bond deviation in period $t+1$ and the utility cost of a wage deviation in period $t+1$. The lower-right block of the matrix H_0 determines the utility loss in the case of a sub-optimal consumption mix. The loss is decreasing in the elasticity of substitution between consumption goods, θ, and depends on the number of consumption goods, I. Finally, conditions (60)-(62) ensure that, in the expression for the expected discounted sum of period utility, after the log-quadratic approximation to expected lifetime utility, one can change the order of integration and summation and all infinite sums converge.

From Proposition 2 one can already see how some parameters will matter for the optimal allocation of attention by a household that cannot attend perfectly to all available information. For example, consider the role of γ. Increasing γ raises the utility loss caused by a deviation of real bond holdings from optimal bond holdings. At the same time, increasing γ

lowers the response of optimal bond holdings to the real interest rate. The relative strength of these two effects will determine whether for a household with a higher γ it is more or less important to be aware of movements in the real interest rate.

6 Aggregation

In this section, we describe issues related to aggregation.

In the following, we will work with log-linearized equations for all aggregate variables. Log-linearizing the equations for aggregate output (8), for aggregate composite consumption (13) and for aggregate composite labor input (14) yields

$$y_t = \frac{1}{I}\sum_{i=1}^{I}\left(\hat{p}_{it} + y_{it}\right), \tag{70}$$

$$c_t = \frac{1}{J}\sum_{j=1}^{J} c_{jt}, \tag{71}$$

and

$$l_t = \frac{1}{I}\sum_{i=1}^{I} l_{it}. \tag{72}$$

Log-linearizing the equations for the price index (29) and for the wage index (50) yields

$$0 = \sum_{i=1}^{I} \hat{p}_{it}, \tag{73}$$

and

$$0 = \sum_{j=1}^{J} \hat{w}_{jt}. \tag{74}$$

Note that the last two equations can also be stated as

$$p_t = \frac{1}{I}\sum_{i=1}^{I} p_{it}, \tag{75}$$

and

$$\tilde{w}_t = \frac{1}{J}\sum_{j=1}^{J} \tilde{w}_{jt}. \tag{76}$$

Furthermore, we will work with log-linearized equations when we aggregate the demands for a particular consumption good or for a particular type of labor. Formally,

$$c_{it} = \frac{1}{J}\sum_{j=1}^{J} c_{ijt}, \tag{77}$$

and

$$l_{jt} = \frac{1}{I} \sum_{i=1}^{I} l_{ijt}. \tag{78}$$

Note that the production function (4) and the monetary policy rule (7) are already log-linear:

$$y_{it} = a_t + a_{it} + \alpha l_{it}, \tag{79}$$

and

$$r_t = \rho_R r_{t-1} + (1 - \rho_R) \left(\phi_\pi \pi_t + \phi_y y_t \right) + \varepsilon_t^R. \tag{80}$$

7 Case 1: Perfect information

In this section, we present the solution of the model under perfect information. This solution will serve as a benchmark. We define the solution of the model under perfect information as follows: In each period t, all agents know the entire history of the economy up to and including period t; firms choose the profit-maximizing price and factor mix; households choose the utility-maximizing consumption vector and nominal wage rate; the government sets the nominal interest rate according to the monetary policy rule, pays subsidies so as to correct the distortions due to market power and chooses a fiscal policy that satisfies the government budget constraint; aggregate variables are given by their respective equations; and households have rational expectations.

The following proposition characterizes real variables at the solution of the model under perfect information after the log-quadratic approximation to the real profit function (see Section 4), the log-quadratic approximation to the expected discounted sum of period utility (see Section 5) and the log-linearization of the equations for the aggregate variables (see Section 6).

Proposition 3 *(Solution of the model under perfect information) A solution to the system of equations (47)-(48), (66)-(69), (70)-(80), (12) and $y_{it} = c_{it}$ with the same initial bond holdings for each household and a non-explosive bond sequence for each household (i.e.*

$\lim_{s\to\infty} E_t \left[\beta^{s+1}\left(\tilde{b}_{j,t+s+1} - \tilde{b}_{j,t+s}\right)\right] = 0$) *satisfies:*

$$y_t = c_t = \frac{1+\psi}{1-\alpha+\alpha\gamma+\psi} a_t, \tag{81}$$

$$l_t = \frac{1-\gamma}{1-\alpha+\alpha\gamma+\psi} a_t, \tag{82}$$

$$\tilde{w}_t = \frac{\gamma+\psi}{1-\alpha+\alpha\gamma+\psi} a_t, \tag{83}$$

$$r_t - E_t\left[\pi_{t+1}\right] = \gamma \frac{1+\psi}{1-\alpha+\alpha\gamma+\psi} E_t\left[a_{t+1} - a_t\right], \tag{84}$$

and

$$\hat{c}_{ijt} = -\theta \hat{p}_{it}, \tag{85}$$

$$\hat{p}_{it} = -\frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\theta} a_{it}, \tag{86}$$

$$\hat{l}_{ijt} = -\eta \hat{w}_{jt}, \tag{87}$$

$$\hat{w}_{jt} = 0. \tag{88}$$

Proof. *See Appendix C.* ■

Under perfect information, aggregate output, aggregate composite consumption, the aggregate composite labor input, the real wage index, and the real interest rate depend only on aggregate technology. The relative price of good i and relative consumption of good i by household j depend only on firm-specific productivity of firm i. The relative wage rate for type j labor and firm i's relative input of type j labor are constant.

Under perfect information, monetary policy has no real effects in this model. Monetary policy does affect nominal variables. The nominal interest rate and inflation follow from the monetary policy rule (80) and the real interest rate (84). 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.[7]

[7] See Woodford (2003), Chapter 2, Proposition 2.8.

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

In this section, we solve the model assuming rational inattention by decision-makers in firms. By rational inattention we mean that decision-makers have limited attention and that they decide how to allocate their attention. For the moment, we continue to assume that households have perfect information to isolate the implications of rational inattention by decision-makers in firms.

8.1 The firms' attention problem

Following Sims (2003), we model attention as an information flow and we model limited attention as a constraint on information flow. To take decisions that are on average close to the profit-maximizing decisions, the decision-maker in a firm has to be aware of changes in the economic environment that cause changes in the profit-maximizing decisions. Being aware of stochastic changes in the environment requires information flow. Decision-makers that have limited attention therefore face a trade-off: Tracking closely particular changes in the environment improves decision making but also uses up valuable information flow.

We formalize this trade-off by letting the decision-maker choose directly the stochastic process for the decision vector, subject to a constraint on information flow. For example, the decision-maker could decide to respond swiftly and correctly with his/her price setting decision to changes in firm-specific productivity but this would require allocating attention to firm-specific productivity. We assume that the decision-maker in a firm chooses the level and the allocation of information flow so as to maximize the expected discounted sum of profits net of the cost of information flow.

Formally, the attention problem of the decision-maker in firm i reads:

$$\max_{\kappa,B_1(L),\ldots,B_3(L),C_1(L),\ldots,C_3(L),\zeta,\chi} \left\{ \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left[\frac{1}{2}\left(x_t - x_t^*\right)' H \left(x_t - x_t^*\right)\right] - \frac{\mu}{1-\beta}\kappa \right\}, \quad (89)$$

where

$$x_t - x_t^* = \begin{pmatrix} p_{it} \\ \hat{l}_{i1t} \\ \vdots \\ \hat{l}_{i(J-1)t} \end{pmatrix} - \begin{pmatrix} p_{it}^* \\ \hat{l}_{i1t}^* \\ \vdots \\ \hat{l}_{i(J-1)t}^* \end{pmatrix}, \tag{90}$$

subject to the equations characterizing the profit-maximizing decisions

$$p_{it}^* = \underbrace{A_1(L)\,\varepsilon_t^A}_{p_{it}^{A*}} + \underbrace{A_2(L)\,\varepsilon_t^R}_{p_{it}^{R*}} + \underbrace{A_3(L)\,\varepsilon_{it}^I}_{p_{it}^{I*}} \tag{91}$$

$$\hat{l}_{ijt}^* = -\eta \hat{w}_{jt}, \tag{92}$$

the equations characterizing the actual decisions

$$p_{it} = \underbrace{B_1(L)\,\varepsilon_t^A + C_1(L)\,\nu_{it}^A}_{p_{it}^A} + \underbrace{B_2(L)\,\varepsilon_t^R + C_2(L)\,\nu_{it}^R}_{p_{it}^R} + \underbrace{B_3(L)\,\varepsilon_{it}^I + C_3(L)\,\nu_{it}^I}_{p_{it}^I} \tag{93}$$

$$\hat{l}_{ijt} = -\zeta\left(\hat{w}_{jt} + \frac{Var(\hat{w}_{jt})}{\chi}\nu_{ijt}^L\right), \tag{94}$$

and the constraint on information flow

$$\mathcal{I}\left(\left\{p_{it}^{A*}, p_{it}^{R*}, p_{it}^{I*}, \hat{l}_{i1t}^*, \ldots, \hat{l}_{i(J-1)t}^*\right\}; \left\{p_{it}^A, p_{it}^R, p_{it}^I, \hat{l}_{i1t}, \ldots, \hat{l}_{i(J-1)t}\right\}\right) \leq \kappa. \tag{95}$$

Here $A_1(L)$ to $A_3(L)$, $B_1(L)$ to $B_3(L)$, and $C_1(L)$ to $C_3(L)$ are infinite-order lag polynomials. The noise terms ν_{it}^A, ν_{it}^R, ν_{it}^I and ν_{ijt}^L appearing in the actual decisions are assumed to follow unit-variance Gaussian white noise processes that are: (i) independent of all other stochastic processes in the economy, (ii) firm-specific, and (iii) independent of each other. The operator $\mathcal{I}$ measures the amount of information that the actual decisions contain about the profit-maximizing decisions. The operator $\mathcal{I}$ is formally defined below.

Proposition 1 states that, after the log-quadratic approximation to the real profit function, the profit-maximizing decisions of firm i in period t are given by equations (47)-(48) and the expected profit loss due to suboptimal decisions is given by equation (45). Objective (89) therefore states that the decision-maker of firm i chooses the level and the allocation of information flow so as to maximize the expected discounted sum of profits net of the cost of information flow.[8] The variable $\kappa \geq 0$ is the information flow devoted to the price setting

[8] A more negative value of expression (45) means a larger expected profit loss due to suboptimal decisions.

and factor mix decisions. The parameter $\mu \geq 0$ is the per-period marginal cost of information flow. This marginal cost of information flow can be interpreted as an opportunity cost (i.e. the cost of devoting less attention to some other activity) or a monetary cost (e.g. an extra wage payment to a manager to improve decision-making).[9]

Equation (91) characterizes the profit-maximizing price setting decision. Here we guess that the profit-maximizing price given by equation (47) has the representation (91) after using equations (71), (76) and $\hat{p}_{it} = p_{it} - p_t$ and after substituting in the equilibrium processes for p_t, c_t, $\tilde{w}_t$, a_t and a_{it}. We will verify this guess. Equation (92) characterizes the profit-maximizing factor mix decision. Here we have simply rewritten the equation for the profit-maximizing factor mix (48) using equations (76) and $\hat{w}_{jt} = \tilde{w}_{jt} - \tilde{w}_t$.

Equation (93) characterizes the actual price setting decision. By choosing the lag polynomials $B_1(L)$ and $C_1(L)$ to $B_3(L)$ and $C_3(L)$, the decision-maker chooses the joint distribution of the profit-maximizing price and the actual price. For example, if $B_1(L) = A_1(L)$ and $C_1(L) = 0$, the price set by the decision-maker responds perfectly to aggregate technology shocks. Similarly, if $B_2(L) = A_2(L)$ and $C_2(L) = 0$, the price set by the decision-maker responds perfectly to monetary policy shocks.

Equation (94) characterizes the actual factor mix decision. By choosing the coefficients ζ and the signal-to-noise ratio χ, the decision-maker chooses the joint distribution of the profit-maximizing factor mix and the actual factor mix. The fact that the decision-maker can only choose two coefficients in equation (94) may seem restrictive compared to equation (93), but we will show below that the firm cannot do better with a less restrictive choice in equation (94).

The information flow constraint (95) restricts the amount of information that the actual decisions contain about the profit-maximizing decisions. We follow Sims (2003) and a large literature in information theory by quantifying information by reduction in uncertainty, where uncertainty is measured by entropy. Let $H(X)$ denote the entropy of the random vector $X = (X_1, \ldots, X_N)$. Entropy is a measure of uncertainty. Let $H(X|Y)$ denote the conditional entropy of the random vector $X = (X_1, \ldots, X_N)$ given knowledge of $Y = (Y_1, \ldots, Y_N)$. Conditional entropy is a measure of conditional uncertainty. The reduction

[9] In equation (90), we use the fact that $\hat{p}_{it} - \hat{p}^*_{it} = p_{it} - p^*_{it}$.

in uncertainty $H(X) - H(X|Y)$ is a measure of the amount of information that Y contains about X. The operator $\mathcal{I}$ in the information flow constraint (95) is defined as

$$\mathcal{I}(\{X_t\};\{Y_t\}) = \lim_{T\to\infty} \frac{1}{T+1} \left[H(X_0,\ldots,X_T) - H(X_0,\ldots,X_T|Y_0,\ldots,Y_T) \right]. \tag{96}$$

The operator $\mathcal{I}$ measures the average per-period amount of information that the stochastic process $\{Y_t\}_{t=0}^{\infty}$ contains about the stochastic process $\{X_t\}_{t=0}^{\infty}$. Thus, the information flow constraint (95) states that the average per-period amount of information that the actual decisions contain about the profit-maximizing decisions cannot exceed the value of κ.

Note that we have assumed that the actual decisions follow a Gaussian process. One can show that a Gaussian process for the actual decisions is optimal because the objective (89) is quadratic and the profit-maximizing decisions (91)-(92) follow a Gaussian process.[10] We have also assumed that the noise appearing in the actual decisions is firm-specific. This assumption accords well with the idea that the friction is the decision-maker's limited attention rather than the availability of information. Finally, we have assumed that the noise terms ν_{it}^A, ν_{it}^R, ν_{it}^I and ν_{ijt}^L are independent of each other. This assumption captures the idea that paying attention to the state of aggregate technology, paying attention to monetary policy disturbances, paying attention to firm-specific productivity and paying attention to relative wage rates are independent activities. Relaxing this assumption is work in progress.

Two remarks are in place before we present solutions of the problem (89)-(95). First, when we solve the problem (89)-(95) numerically, we turn this infinite-dimensional problem into a finite-dimensional problem by parameterizing each infinite-order lag polynomial $B_1(L)$ to $B_3(L)$ and $C_1(L)$ to $C_3(L)$ as a lag-polynomial of an ARMA(p,q) process where p and q are finite. Second, when a variable appearing in the information flow constraint (95) is (or may be) non-stationary, we replace the original variable by its first difference in the information flow constraint to ensure that entropy is always well defined.

[10]See Sims (2006) or Section VIIA in Maćkowiak and Wiederholt (2009).

8.2 Computing the equilibrium of the model

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 (91) and a guess concerning the process for the relative wage rate in equation (92). Second, we solve the firms' attention problem (89)-(95). Third, we aggregate the individual prices to obtain the aggregate price level:

$$p_t = \frac{1}{I} \sum_{i=1}^{I} p_{it}. \tag{97}$$

Fourth, we compute the aggregate dynamics implied by those price level dynamics. Recall that in this section we assume that households have perfect information. The households' optimality conditions (66)-(68), equations (70)-(80), equation (12), $y_{it} = c_{it}$ and the assumption that aggregate technology follows a first-order autoregressive process imply that the following equations have to be satisfied in equilibrium:

$$c_t = E_t \left[-\frac{1}{\gamma} \left(r_t - p_{t+1} + p_t \right) + c_{t+1} \right], \tag{98}$$

$$\tilde{w}_t = \gamma c_t + \psi l_t, \tag{99}$$

$$y_t = c_t, \tag{100}$$

$$y_t = a_t + \alpha l_t, \tag{101}$$

$$a_t = \rho_A a_{t-1} + \varepsilon_t^A, \tag{102}$$

$$r_t = \rho_R r_{t-1} + (1 - \rho_R) \left[\phi_\pi \left(p_t - p_{t-1} \right) + \phi_y y_t \right] + \varepsilon_t^R. \tag{103}$$

Here E_t denotes the expectation operator conditioned on the entire history of the economy up to and including period t. 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 $(c_t, \tilde{w}_t, y_t, l_t, a_t, r_t)$ implied by the price level dynamics. Fifth, we compute the law of motion for the profit-maximizing price. The firms' optimality condition (47) and equations (71), (76) and $\hat{p}_{it} = p_{it} - p_t$ imply that the profit-maximizing price is given by

$$p_{it}^* = p_t + \frac{\frac{1-\alpha}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\vartheta} c_t + \frac{1}{1 + \frac{1-\alpha}{\alpha}\vartheta} \tilde{w}_t - \frac{\frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\vartheta} \left(a_t + a_{it} \right). \tag{104}$$

By substituting the law of motion for p_t, c_t, $\tilde{w}_t$, a_t and a_{it} into the last equation, we obtain the law of motion for the profit-maximizing price. In the last equation, we set $\vartheta = \theta$ because the households' optimality condition (68) and equations (71), (73) and (77) imply that the demand function for good i has the form (28)-(30) with a price elasticity of demand $\vartheta = \theta$. If the process for the profit-maximizing price differs from our guess, we update the guess until a fixed point is reached.

Finally, we derive the equilibrium relative wage rates and the equilibrium factor mix. Suppose that firms choose a value for ζ that exceeds 1 and a value for χ that is strictly positive. Then, each firm can attain the profit-maximizing factor mix without any information flow. Thus, no firm has an incentive to deviate. The argument is the following. Equation (94) and equations (72) and (78) imply that the labor demand function for each type of labor has the form (49)-(51). Since all households have exactly the same decision problem, all households set the same wage rate. It follows from equation (76) that $\tilde{w}_t = \tilde{w}_{jt}$, or equivalently $\hat{w}_{jt} = 0$. Thus, in equilibrium the profit-maximizing factor mix is constant ($\hat{l}^*_{ijt} = 0$), implying that each firm can attain the profit-maximizing factor mix ($\hat{l}_{ijt} = 0$) without any information flow.

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 $\beta = 0.99$, $\gamma = 1$, $\psi = 0$, $\theta = 4$, $\alpha = 2/3$ and $\eta = 4$.

To set the parameters governing the process for aggregate technology, equation (102), 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 (101) 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.[11] We then fit equation (102) 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.

[11] 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.

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.[12] We fit the Taylor rule (103) 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.

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 (2008) 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 (2008). This yields a standard deviation of the innovation to firm-specific productivity equal to 0.22.

We compute the solution of the model by fixing the marginal value of information flow instead of κ. The overall information flow, κ, is then determined within the model. The idea is the following. When the marginal value of information flow is high, decision-makers in firms have a high incentive to increase information flow in order to take better decisions. In contrast, when the marginal value of information flow is low, decision-makers in firms have little incentive to increase information flow. We set the marginal value of information flow equal to 0.1 percent of a firm's steady state output. We obtain this marginal value of information flow in equilibrium by setting the marginal cost of information flow in objective (89) to $\mu = (0.001)\, Y_i$.

We first report the optimal allocation of attention at the rational inattention fixed point. The decision-maker in a firm allocates 3.1 bits of information flow to tracking firm-specific productivity, 1 bit of information flow to tracking aggregate technology, and 0.35 bits of

[12] 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.

information flow to tracking monetary policy. The expected per-period loss in profits due to imperfect tracking of firm-specific productivity equals 0.07 percent of the firm's steady state output; the expected per-period loss in profits due to imperfect tracking of aggregate technology equals 0.05 percent of the firm's steady state output; and the expected per-period loss in profits due to imperfect tracking of monetary policy equals 0.03 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 0.15 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 equilibrium under perfect information derived in Section 7 (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 then 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. Decision-makers in firms decide to allocate about three times as much attention to aggregate technology than to monetary policy. Therefore, prices respond faster to aggregate technology shocks than to monetary policy shocks. As a result, the output gap is negative for only 5 quarters after an aggregate technology shock, while the output gap is negative for 10 quarters after a monetary policy shock.[13]

Figure 3 shows the impulse response of an individual price to a firm-specific productivity shock. Prices respond almost perfectly to firm-specific productivity shocks. The reason is the optimal allocation of attention. Decision-makers in firms decide to pay close attention to firm-specific productivity.

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. 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 Calvo model with perfect information (red lines with crosses). The impulse responses to a monetary policy shock are essentially identical in the two models, while the impulse responses to an aggregate technology shock are quite different in the two models. In particular, 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. This is because decision-makers in firms in the benchmark economy decide to allocate about three times as much attention to aggregate technology than to monetary policy.

In the benchmark economy and in the Calvo model, firms experience profit losses due to deviations of the actual price from the profit-maximizing price. In the benchmark economy,

[13] See also Paciello (2008). 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.

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. Specifically, the expected loss in profits due to sub-optimal price responses to aggregate conditions is 23 times smaller than in the Calvo model. In addition, the expected loss in profits due to sub-optimal price responses to firm-specific conditions is 57 times smaller than in the Calvo model. The main reason for this result is the optimal allocation of attention. In the benchmark economy, prices respond slowly to monetary policy shocks, but fairly quickly to aggregate technology shocks, and very quickly to micro-level shocks. In contrast, in the Calvo model, prices respond slowly to all those shocks. Another reason for this result 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.

9 Case 3: Firms and households have limited attention

We now study the implications of adding rational inattention by households. We first make two simplifying assumptions to focus on the implications of rational inattention by households for consumption behavior. In particular, we assume that households set real wage rates (instead of nominal wage rates) and $\psi = 0$. One can show analytically that these two assumptions imply that the optimal wage setting behavior under both perfect information and limited attention satisfies

$$\tilde{w}_{jt} = \gamma c_{jt}. \tag{105}$$

The reason is as follows. When households set real wage rates and $\psi = 0$, each household only needs to know his/her own consumption to be on the labor supply curve. Knowing own consumption does not require any information flow. Hence, the assumptions that households set real wage rates (instead of nominal wage rates) and $\psi = 0$ allow us to study in isolation the implications of rational inattention by households for consumption behavior.

9.1 The households' attention problem

The attention problem of household j reads:

$$\max_{\kappa,B(L),C(L),\vartheta,\chi} \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[\frac{1}{2} \left(x_t - x_t^*\right)' H_0 \left(x_t - x_t^*\right) + \left(x_t - x_t^*\right)' H_1 \left(x_{t+1} - x_{t+1}^*\right) \right] - \frac{\mu}{1-\beta}\kappa, \tag{106}$$

where

$$x_t - x_t^* = \begin{pmatrix} \tilde{b}_{jt} \\ \tilde{w}_{jt} \\ \hat{c}_{1jt} \\ \vdots \\ \hat{c}_{I-1jt} \end{pmatrix} - \begin{pmatrix} \tilde{b}_{jt}^* \\ \tilde{w}_{jt}^* \\ \hat{c}_{1jt}^* \\ \vdots \\ \hat{c}_{I-1jt}^* \end{pmatrix}, \tag{107}$$

subject to an equation linking an argument of the objective and two decision variables

$$\tilde{b}_{jt} - \tilde{b}_{jt}^* = -\frac{1}{\omega_B} \sum_{s=0}^{t} \left(\frac{1}{\beta}\right)^{t-s} \left[\left(c_{js} - c_{js}^*\right) + \zeta \omega_W \left(\tilde{w}_{js} - \tilde{w}_{js}^*\right) \right], \tag{108}$$

the equations characterizing the household's optimal behavior under perfect information

$$c_{jt}^* = \underbrace{A_1(L)\,\varepsilon_t^A}_{c_{jt}^{A*}} + \underbrace{A_2(L)\,\varepsilon_t^R}_{c_{jt}^{R*}} \tag{109}$$

$$\tilde{w}_{jt}^* = \gamma c_{jt}^* \tag{110}$$

$$\hat{c}_{ijt}^* = -\theta \hat{p}_{it}, \tag{111}$$

the equations characterizing the household's actual behavior

$$c_{jt} = \underbrace{B_1(L)\,\varepsilon_t^A + C_1(L)\,\nu_{jt}^A}_{c_{jt}^A} + \underbrace{B_2(L)\,\varepsilon_t^R + C_2(L)\,\nu_{jt}^R}_{c_{jt}^R} \tag{112}$$

$$\tilde{w}_{jt} = \gamma c_{jt} \tag{113}$$

$$\hat{c}_{ijt} = -\vartheta \left(\hat{p}_{it} + \frac{Var(\hat{p}_{it})}{\chi} \nu_{ijt}^I \right), \tag{114}$$

and the information flow constraint

$$\mathcal{I}\left(\left\{ c_{jt}^{A*}, c_{jt}^{R*}, \hat{c}_{1jt}^*, \ldots, \hat{c}_{I-1jt}^* \right\}; \left\{ c_{jt}^A, c_{jt}^R, \hat{c}_{1jt}, \ldots, \hat{c}_{I-1jt} \right\} \right) \leq \kappa. \tag{115}$$

Here $A_1(L)$, $A_2(L)$, $B_1(L)$, $B_2(L)$, $C_1(L)$ and $C_2(L)$ are infinite-order lag polynomials. The noise terms ν_{jt}^A, ν_{jt}^R and ν_{ijt}^I in the actual decisions are assumed to follow unit-variance

Gaussian white noise processes that are: (i) independent of all other stochastic processes in the economy, (ii) household-specific, and (iii) independent of each other. The operator $\mathcal{I}$ measures the amount of information that the household's actual decisions contain about the household's optimal decisions under perfect information.

In equations (112)-(114), we assume that the household chooses a consumption vector and a real wage rate. The household's real bond holdings then follow from equation (108), which follows from the flow budget constraint (69).

Finally, we assume that, in period -1, the economy is in the non-stochastic steady state and all households know that the economy is in the non-stochastic steady state.

When we solve the problem (106)-(115) numerically, we turn this infinite-dimensional problem into a finite-dimensional problem by parameterizing each infinite-order lag polynomial $B_1(L)$, $B_2(L)$, $C_1(L)$ and $C_2(L)$ as a lag-polynomial of an ARMA(p,q) process where p and q are finite.

9.2 Benchmark parameter values and solution

We assume the same parameter values as in the benchmark economy in Section 8.3. We have to choose values for three additional parameters: ω_B, ω_W and the household's marginal value of information flow. We set $\omega_B = 4$ and $\omega_W = 0.95$. We set the household's marginal value of information flow equal to 0.1 percent of the household's steady state composite consumption. We obtain this marginal value of information flow by setting the marginal cost of information flow in objective (106) to $\mu = (0.001)\, C_j$.

We begin with the following experiment in order to get a first idea of how rational inattention by households affects consumption behavior. We study the optimal allocation of attention by an individual household assuming that decision-makers in firms have limited attention and all other households have perfect information, i.e. we study the optimal allocation of attention by an individual household at the fixed point derived in Section 8. The optimal allocation of attention by the household has the following features. The household allocates 0.31 bits of information flow to tracking aggregate technology and 0.12 bits of information flow to tracking monetary policy. The expected per-period loss in utility due to imperfect tracking of aggregate technology equals 0.02 percent of the household's steady state composite consumption, and the expected per-period loss in utility due to imperfect tracking of monetary policy also equals 0.02 percent of the household's steady state composite consumption. Figure 6 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 green line with circles is the impulse response under perfect information, while the black line with diamonds is the impulse response under limited attention. We would like to point out four results. First, there are sizeable differences between the impulse responses of consumption under perfect information and the impulse responses of 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.

Next, we solve for the fixed point when decision-makers in firms and all households have limited attention. When we add rational inattention by households, the decision-maker in a firm allocates 1 bit of information flow to tracking aggregate technology and 0.3 bits of information flow to tracking monetary policy. Less attention gets allocated to tracking monetary policy compared with the case when households had perfect information. The expected per-period loss in profits due to imperfect tracking of aggregate technology is approximately unaffected. The expected per-period loss in profits due to imperfect tracking of monetary policy falls to 0.02 percent of the firm's steady state output. Each household allocates 0.34 bits of information flow to tracking aggregate technology and 0.18 bits of information flow to tracking monetary policy. The expected per-period loss in utility due to imperfect tracking of aggregate technology equals 0.03 percent of the household's steady state composite consumption. The expected per-period loss in utility due to imperfect tracking of monetary policy also equals 0.03 percent of the household's steady state composite consumption.

Figures 7 and 8 show equilibrium impulse responses of the price level, inflation, consumption, and the nominal interest rate (black lines with asterisks). For comparison, the figures also include impulse responses of the same variables at the fixed point when decision-makers in firms have limited attention and households have perfect information (green lines with circles). When we add rational inattention by households, the impulse responses to a monetary policy shock change considerably, despite the fact that the utility loss from sub-optimal behavior is very small and the marginal value of information flow is very low. The response of the price level to a monetary policy shock becomes more dampened. The response of inflation to a monetary policy shock becomes more persistent. The response of consumption to a monetary policy shock becomes hump-shaped. See Figure 7. Adding rational inattention by households has only a small effect on the impulse responses to an aggregate technology shock. See Figure 8. One reason is that households allocate twice more attention to tracking aggregate technology than to tracking monetary policy.

10 Conclusion

We have solved a dynamic stochastic general equilibrium model in which decision-makers in firms and households have limited attention and decide how to allocate their attention. In contrast to the existing literature on rational inattention, we solve a dynamic stochastic general equilibrium model. In contrast to the existing literature on business cycle models with imperfect information, the information structure is the outcome of an optimization problem.

The impulse responses of prices under rational inattention by decision-makers in firms have several properties of empirical impulse responses: (i) the price level responds slowly to monetary policy shocks, (ii) the price level responds faster to aggregate technology shocks, and (iii) prices respond very fast to disaggregate shocks.[14] These impulse responses imply that 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.

The impulse response of consumption to a monetary policy shock under rational inattention by households looks similar to the impulse response of consumption to a monetary policy disturbance in a model with habit formation.

These results suggest that the slow responses of prices and consumption to monetary policy shocks that are usually modeled with a Calvo price-setting friction and habit formation may have a different origin.

[14] For empirical evidence on the response of the price level to aggregate technology shocks, see Altig, Christiano, Eichenbaum, and Linde (2004). For empirical evidence on the response of prices to disaggregate shocks, see Boivin, Giannoni, and Mihov (2009) and Maćkowiak, Moench, and Wiederholt (2009).

A Proof of Proposition 1

First, we introduce notation. Let x_t denote the vector of all variables appearing in the real profit function f that the firm can affect

$$x_t' = \begin{pmatrix} \hat{p}_{it} & \hat{l}_{i1t} & \cdots & \hat{l}_{i(J-1)t} \end{pmatrix}. \tag{116}$$

Let z_t denote the vector of all variables appearing in the real profit function f that the firm takes as given

$$z_t' = \begin{pmatrix} a_t & a_{it} & c_{1t} & \cdots & c_{Jt} & \tilde{w}_{1t} & \cdots & \tilde{w}_{Jt} \end{pmatrix}. \tag{117}$$

Second, we compute a quadratic approximation to the expected discounted sum of profits (39) at the non-stochastic steady state. Let $\tilde{f}$ denote the second-order Taylor approximation to f at the non-stochastic steady state. We have

$$\begin{aligned} & E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t \tilde{f}\left(x_t, z_t\right)\right] \\ = \;& E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t \left(f\left(0,0\right) + h_x' x_t + h_z' z_t + \frac{1}{2}x_t' H_x x_t + x_t' H_{xz} z_t + \frac{1}{2} z_t' H_z z_t\right)\right], \end{aligned} \tag{118}$$

where h_x is the vector of first derivatives of f with respect to x_t evaluated at the non-stochastic steady state, h_z is the vector of first derivatives of f with respect to z_t evaluated at the non-stochastic steady state, H_x is the matrix of second derivatives of f with respect to x_t evaluated at the non-stochastic steady state, H_z is the matrix of second derivatives of f with respect to z_t evaluated at the non-stochastic steady state, and H_{xz} is the matrix of second derivatives of f with respect to x_t and z_t evaluated at the non-stochastic steady state. Third, we rewrite equation (118) using condition (44). Let v_t denote the following vector

$$v_t' = \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}, \tag{119}$$

and let $v_{m,t}$ denote the mth element of v_t. Condition (44) implies that

$$\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left|f\left(0,0\right) + h_x' x_t + h_z' z_t + \frac{1}{2}x_t' H_x x_t + x_t' H_{xz} z_t + \frac{1}{2} z_t' H_z z_t\right| < \infty. \tag{120}$$

It follows that one can rewrite equation (118) as

$$\begin{aligned} & E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t\tilde{f}(x_t,z_t)\right] \\ = & \sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[f(0,0)+h_x'x_t+h_z'z_t+\frac{1}{2}x_t'H_x x_t+x_t'H_{xz}z_t+\frac{1}{2}z_t'H_z z_t\right]. \end{aligned} \tag{121}$$

See Rao (1973), p. 111. Condition (44) also implies that the infinite sum on the right-hand side of equation (121) 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

$$h_x + H_x x_t^* + H_{xz} z_t = 0. \tag{122}$$

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

$$x_t^* = -H_x^{-1} h_x - H_x^{-1} H_{xz} z_t. \tag{123}$$

Hence, x_t^* is uniquely determined and the vector v_t with $x_t = x_t^*$ satisfies condition (44). Fifth, equation (121) implies that

$$\begin{aligned} & E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t\tilde{f}(x_t,z_t)\right]-E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t\tilde{f}(x_t^*,z_t)\right] \\ = & \sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[h_x'(x_t-x_t^*)+\frac{1}{2}x_t'H_x x_t-\frac{1}{2}x_t^{*\prime}H_x x_t^*+(x_t-x_t^*)'H_{xz}z_t\right]. \end{aligned} \tag{124}$$

Using equation (122) to substitute for $H_{xz}z_t$ in the last equation and rearranging yields

$$\begin{aligned} & E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t\tilde{f}(x_t,z_t)\right]-E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t\tilde{f}(x_t^*,z_t)\right] \\ = & \sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}(x_t-x_t^*)'H_x(x_t-x_t^*)\right]. \end{aligned} \tag{125}$$

Sixth, we compute the vector of first derivatives and the matrices of second derivatives appearing in equations (123) and (125). We obtain

$$h_x = 0, \tag{126}$$

$$H_x = -C_j^{-\gamma}\tilde{W}L_i \begin{bmatrix} \frac{\vartheta}{\alpha}\left(1+\frac{1-\alpha}{\alpha}\vartheta\right) & 0 & \cdots & \cdots & 0 \\ 0 & \frac{2}{\eta J} & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J} \\ \vdots & \frac{1}{\eta J} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1}{\eta J} \\ 0 & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J} & \frac{2}{\eta J} \end{bmatrix}, \tag{127}$$

and

$$H_{xz} = C_j^{-\gamma}\tilde{W}L_i \begin{bmatrix} -\frac{\vartheta}{\alpha}\frac{1}{\alpha} & -\frac{\vartheta}{\alpha}\frac{1}{\alpha} & \frac{\vartheta}{\alpha}\frac{1-\alpha}{\alpha}\frac{1}{J} & \cdots & \frac{\vartheta}{\alpha}\frac{1-\alpha}{\alpha}\frac{1}{J} & \frac{\vartheta}{\alpha}\frac{1}{J} & \cdots & \cdots & \frac{\vartheta}{\alpha}\frac{1}{J} & \frac{\vartheta}{\alpha}\frac{1}{J} \\ 0 & 0 & 0 & \cdots & 0 & -\frac{1}{J} & 0 & \cdots & 0 & \frac{1}{J} \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & -\frac{1}{J} & \frac{1}{J} \end{bmatrix}, \tag{128}$$

where we used equation (36) in equations (127)-(128). Seventh, substituting equations (126)-(128) into equation (122) yields the following system of J equations:

$$\hat{p}_{it}^{*} = \frac{\frac{1-\alpha}{\alpha}}{1+\frac{1-\alpha}{\alpha}\vartheta}\left(\frac{1}{J}\sum_{j=1}^{J} c_{jt}\right) + \frac{1}{1+\frac{1-\alpha}{\alpha}\vartheta}\left(\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}\right) - \frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\vartheta}\left(a_t + a_{it}\right), \tag{129}$$

and

$$\forall j \neq J : \hat{l}_{ijt}^{*} + \sum_{k=1}^{J-1}\hat{l}_{ikt}^{*} = -\eta\left(\tilde{w}_{jt} - \tilde{w}_{Jt}\right). \tag{130}$$

Finally, we rewrite equation (130). Summing equation (130) over all $j \neq J$ yields

$$\sum_{j=1}^{J-1}\hat{l}_{ijt}^{*} = -\eta\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt} + \eta\tilde{w}_{Jt}. \tag{131}$$

Substituting the last equation back into equation (130) yields

$$\forall j \neq J : \hat{l}_{ijt}^{*} = -\eta\left(\tilde{w}_{jt} - \frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}\right). \tag{132}$$

Collecting equations (125), (127), (129) and (132), we arrive at Proposition 1.

B Proof of Proposition 2

First, we introduce notation. In each period $t \geq 0$, let x_t denote the vector of all variables appearing in the period utility function (55) that the household can affect in period t

$$x_t' = \begin{pmatrix} \tilde{b}_{jt} & \tilde{w}_{jt} & \hat{c}_{1jt} & \cdots & \hat{c}_{I-1jt} \end{pmatrix}, \tag{133}$$

and, in each period $t \geq 0$, let z_t denote the vector of all variables appearing in the period utility function (55) that the household takes as given

$$z_t' = \begin{pmatrix} r_{t-1} & \pi_t & \tilde{w}_t & l_t & \tilde{d}_t & \tilde{t}_t & \hat{p}_{1t} & \cdots & \hat{p}_{It} \end{pmatrix}. \tag{134}$$

There is one variable appearing in the period utility function (55) that is neither an element of x_t nor an element of z_t: the predetermined variable $\tilde{b}_{jt-1}$. For ease of exposition, we define the $(1+I)$-dimensional column vector x_{-1} by

$$x_{-1}' = \begin{pmatrix} \tilde{b}_{j,-1} & 0 & \cdots & 0 \end{pmatrix}, \tag{135}$$

because then, in each period $t \geq 0$, the predetermined variable $\tilde{b}_{jt-1}$ is an element of x_{t-1}. Let g denote the functional that is obtained by multiplying the period utility function (55) by β^t and summing over all t from zero to infinity. Let $\tilde{g}$ denote the second-order Taylor approximation to g at the non-stochastic steady state. Finally, let $E_{j,-1}$ denote the expectation operator conditioned on information of household j in period -1. Second, we compute a log-quadratic approximation to the expected discounted sum of period utility around the non-stochastic steady state. We obtain

$$\begin{aligned} & E_{j,-1}\left[\tilde{g}\left(x_{-1}, x_0, z_0, x_1, z_1, x_2, z_2, \ldots\right)\right] \\ = \; & E_{j,-1}\left[\begin{array}{c} g\left(0,0,0,0,0,0,0,\ldots\right) \\ +\sum_{t=0}^{\infty}\beta^t\begin{pmatrix} h_x' x_t + h_z' z_t \\ +\frac{1}{2}x_t' H_{x,-1} x_{t-1} + \frac{1}{2}x_t' H_{x,0} x_t + \frac{1}{2}x_t' H_{x,1} x_{t+1} \\ +\frac{1}{2}x_t' H_{xz,0} z_t + \frac{1}{2}x_t' H_{xz,1} z_{t+1} \\ +\frac{1}{2}z_t' H_{z,0} z_t + \frac{1}{2}z_t' H_{zx,-1} x_{t-1} + \frac{1}{2}z_t' H_{zx,0} x_t \end{pmatrix} \\ +\beta^{-1}\left(h_{-1}' x_{-1} + \frac{1}{2}x_{-1}' H_{-1} x_{-1} + \frac{1}{2}x_{-1}' H_{x,1} x_0 + \frac{1}{2}x_{-1}' H_{xz,1} z_0\right) \end{array}\right], \end{aligned} \tag{136}$$

where $\left(\beta^t h_x\right)$ is the vector of first derivatives of g with respect to x_t evaluated at the non-stochastic steady state, $\left(\beta^t h_z\right)$ is the vector of first derivatives of g with respect to z_t

evaluated at the non-stochastic steady state, $\left(\beta^t H_{x,\tau}\right)$ is the matrix of second derivatives of g with respect to x_t and $x_{t+\tau}$ evaluated at the non-stochastic steady state, $\left(\beta^t H_{z,\tau}\right)$ is the matrix of second derivatives of g with respect to z_t and $z_{t+\tau}$ evaluated at the non-stochastic steady state, $\left(\beta^t H_{xz,\tau}\right)$ is the matrix of second derivatives of g with respect to x_t and $z_{t+\tau}$ evaluated at the non-stochastic steady state, and $\left(\beta^t H_{zx,\tau}\right)$ is the matrix of second derivatives of g with respect to z_t and $x_{t+\tau}$ evaluated at the non-stochastic steady state. Finally, $\left(\beta^{-1} h_{-1}\right)$ is a $(1+I)$-dimensional column vector whose first element equals the first derivative of g with respect to $\tilde{b}_{j,-1}$ evaluated at the non-stochastic steady state and $\left(\beta^{-1} H_{-1}\right)$ is a $(1+I) \times (1+I)$ matrix whose upper left element equals the second derivative of g with respect to $\tilde{b}_{j,-1}$ evaluated at the non-stochastic steady state. Note that only certain quadratic terms appear on the right-hand side of equation (136) because: (i) for all $t \geq 0$, the vector of first derivatives of g with respect to x_t depends only on elements of x_{t-1}, x_t, x_{t+1}, z_t and z_{t+1}, (ii) for all $t \geq 0$, the vector of first derivatives of g with respect to z_t depends only on elements of z_t, x_{t-1} and x_t, and (iii) the first derivative of g with respect to $\tilde{b}_{j,-1}$ depends only on elements of x_{-1}, x_0 and z_0. Furthermore, note that, when we write the vector of first derivatives of g with respect to x_t evaluated at the non-stochastic steady state as $\left(\beta^t h_x\right)$, we exploit the fact that this vector of first derivatives depends on t only through the multiplicative term β^t. Third, we rewrite equation (136) using conditions (60)-(62). For all $t \geq 0$, let v_t denote the following vector

$$v_t' = \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}. \tag{137}$$

For $t = -1$, let v_t denote a $(8+2I)$-dimensional column vector whose first element equals $\tilde{b}_{j,-1}$ and all other elements equal zero. Let $v_{m,t}$ denote the mth element of v_t. Condition (62) implies that, for all m and n and for $\tau = 0, 1$,

$$\sum_{t=0}^{\infty} \beta^t E_{j,-1} \left| v_{m,t} v_{n,t+\tau} \right| < \infty. \tag{138}$$

Condition (61) implies that condition (138) also holds for $\tau = -1$. It follows that, for all m and n and for $\tau = 0, 1, -1$,

$$E_{j,-1}\left[\sum_{t=0}^{\infty} \beta^t v_{m,t} v_{n,t+\tau}\right] = \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[v_{m,t} v_{n,t+\tau}\right]. \tag{139}$$

See Rao (1973), p. 111. Furthermore, conditions (61)-(62) imply that the infinite sum on the right-hand side of equation (139) converges to an element in $\mathbb{R}$. Thus, conditions (61)-(62) imply that one can rewrite equation (136) as

$$\begin{aligned}
& E_{j,-1}\left[\tilde{g}\left(x_{-1},x_0,z_0,x_1,z_1,x_2,z_2,\ldots\right)\right] \\
= \; & g\left(0,0,0,0,0,0,0,\ldots\right)+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[h_x' x_t\right]+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[h_z' z_t\right] \\
& +\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}x_t' H_{x,-1}x_{t-1}\right]+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}x_t' H_{x,0}x_t\right] \\
& +\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}x_t' H_{x,1}x_{t+1}\right]+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}x_t' H_{xz,0}z_t\right]+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}x_t' H_{xz,1}z_{t+1}\right] \\
& +\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}z_t' H_{z,0}z_t\right]+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}z_t' H_{zx,-1}x_{t-1}\right]+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}z_t' H_{zx,0}x_t\right] \\
& +\beta^{-1}E_{j,-1}\left[h_{-1}' x_{-1}+\frac{1}{2}x_{-1}' H_{-1}x_{-1}+\frac{1}{2}x_{-1}' H_{x,1}x_0+\frac{1}{2}x_{-1}' H_{xz,1}z_0\right], \qquad (140)
\end{aligned}$$

and that each infinite sum on the right-hand side of equation (140) converges to an element in $\mathbb{R}$. In addition, conditions (60)-(61) ensure that the term in the last line on the right-hand side of equation (140) is finite. Finally, using $H_{xz,0}=H_{zx,0}'$, $H_{xz,1}=\beta H_{zx,-1}'$ and $H_{x,1}=\beta H_{x,-1}'$ one can rewrite equation (140) as

$$\begin{aligned}
& E_{j,-1}\left[\tilde{g}\left(x_{-1},x_0,z_0,x_1,z_1,x_2,z_2,\ldots\right)\right] \\
= \; & g\left(0,0,0,0,0,0,0,\ldots\right)+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[h_x' x_t\right]+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[h_z' z_t\right] \\
& +\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}x_t' H_{x,0}x_t\right]+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[x_t' H_{x,1}x_{t+1}\right]+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[x_t' H_{xz,0}z_t\right] \\
& +\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[x_t' H_{xz,1}z_{t+1}\right]+\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}z_t' H_{z,0}z_t\right] \\
& +\beta^{-1}E_{j,-1}\left[h_{-1}' x_{-1}+\frac{1}{2}x_{-1}' H_{-1}x_{-1}+x_{-1}' H_{x,1}x_0+x_{-1}' H_{xz,1}z_0\right]. \qquad (141)
\end{aligned}$$

Fourth, we define the process $\{x_t^*\}$. Let E_t denote the expectation operator conditioned on the entire history of the economy up to and including period t. The process $\{x_t^*\}$ is defined by the following three requirements: (i) x_{-1}^* is given by equation (135), (ii) in each period $t\geq 0$, x_t^* satisfies

$$E_t\left[h_x+H_{x,-1}x_{t-1}^*+H_{x,0}x_t^*+H_{x,1}x_{t+1}^*+H_{xz,0}z_t+H_{xz,1}z_{t+1}\right]=0, \qquad (142)$$

and (iii) the vector v_t with $x_t = x_t^*$ satisfies conditions (60)-(62). Fifth, we derive a result that we will use below. Multiplying equation (142) by $(x_t - x_t^*)'$ and using the fact that E_t is the expectation operator conditioned on the entire history of the economy up to and including period t yields

$$E_t \left[(x_t - x_t^*)' \left(h_x + H_{x,-1} x_{t-1}^* + H_{x,0} x_t^* + H_{x,1} x_{t+1}^* + H_{xz,0} z_t + H_{xz,1} z_{t+1} \right) \right] = 0. \quad (143)$$

Taking the expectation conditioned on information of household j in period $t = -1$ and using the law of iterated expectations yields

$$E_{j,-1} \left[(x_t - x_t^*)' \left(h_x + H_{x,-1} x_{t-1}^* + H_{x,0} x_t^* + H_{x,1} x_{t+1}^* + H_{xz,0} z_t + H_{xz,1} z_{t+1} \right) \right] = 0. \quad (144)$$

Rearranging the last equation yields

$$\begin{aligned} & E_{j,-1} \left[(x_t - x_t^*)' \left(h_x + H_{xz,0} z_t + H_{xz,1} z_{t+1} \right) \right] \\ = \quad & -E_{j,-1} \left[(x_t - x_t^*)' \left(H_{x,-1} x_{t-1}^* + H_{x,0} x_t^* + H_{x,1} x_{t+1}^* \right) \right]. \end{aligned} \quad (145)$$

Sixth, we derive another result that we will use below. By the Cauchy-Schwarz inequality, for each period $t \geq 0$, for $\tau = 0, 1, -1$ and for all m and n,

$$\left(E_{j,-1} \left[x_{m,t} x_{n,t+\tau}^* \right] \right)^2 \leq E_{j,-1} \left[x_{m,t}^2 \right] E_{j,-1} \left[x_{n,t+\tau}^{*2} \right]. \quad (146)$$

Conditions (60) and (62) and the definition of the process $\{x_t^*\}$ therefore imply that there exist two constants $\delta < (1/\beta)$ and $A \in \mathbb{R}$ such that, for each period $t \geq 0$, for $\tau = 0, 1, -1$ and for all m and n,

$$\left| E_{j,-1} \left[x_{m,t} x_{n,t+\tau}^* \right] \right| < \delta^t A. \quad (147)$$

It follows that $\left\{ \sum_{t=0}^{T} \beta^t E_{j,-1} \left[x_{m,t} x_{n,t+\tau}^* \right] \right\}_{T=0}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$, implying that $\sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[x_{m,t} x_{n,t+\tau}^* \right]$ converges to an element in $\mathbb{R}$. Seventh, it follows from equation

(141) that

$$
\begin{aligned}
& E_{j,-1}\left[\tilde{g}\left(x_{-1},x_0,z_0,x_1,z_1,x_2,z_2,\ldots\right)\right]-E_{j,-1}\left[\tilde{g}\left(x_{-1}^*,x_0^*,z_0,x_1^*,z_1,x_2^*,z_2,\ldots\right)\right] \\
= & \sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}x_t'H_{x,0}x_t+x_t'H_{x,1}x_{t+1}-\frac{1}{2}x_t^{*\prime}H_{x,0}x_t^*-x_t^{*\prime}H_{x,1}x_{t+1}^*\right] \\
& +\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\left(x_t-x_t^*\right)'\left(h_x+H_{xz,0}z_t+H_{xz,1}z_{t+1}\right)\right] \\
& +\beta^{-1}E_{j,-1}\left[h_{-1}'x_{-1}+\frac{1}{2}x_{-1}'H_{-1}x_{-1}+x_{-1}'H_{x,1}x_0+x_{-1}'H_{xz,1}z_0\right] \\
& -\beta^{-1}E_{j,-1}\left[h_{-1}'x_{-1}^*+\frac{1}{2}x_{-1}^{*\prime}H_{-1}x_{-1}^*+x_{-1}^{*\prime}H_{x,1}x_0^*+x_{-1}^{*\prime}H_{xz,1}z_0\right]. \qquad (148)
\end{aligned}
$$

Substituting $x_{-1}^*=x_{-1}$ and equation (145) into equation (148) yields

$$
\begin{aligned}
& E_{j,-1}\left[\tilde{g}\left(x_{-1},x_0,z_0,x_1,z_1,x_2,z_2,\ldots\right)\right]-E_{j,-1}\left[\tilde{g}\left(x_{-1}^*,x_0^*,z_0,x_1^*,z_1,x_2^*,z_2,\ldots\right)\right] \\
= & \sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}x_t'H_{x,0}x_t+x_t'H_{x,1}x_{t+1}-\frac{1}{2}x_t^{*\prime}H_{x,0}x_t^*-x_t^{*\prime}H_{x,1}x_{t+1}^*\right] \\
& -\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\left(x_t-x_t^*\right)'\left(H_{x,-1}x_{t-1}^*+H_{x,0}x_t^*+H_{x,1}x_{t+1}^*\right)\right] \\
& +\beta^{-1}E_{j,-1}\left[x_{-1}'H_{x,1}\left(x_0-x_0^*\right)\right].
\end{aligned}
$$

Finally, rearranging the right-hand side of the last equation using that (i) $\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[x_t'H_{x,\tau}x_{t+\tau}^*\right]$ converges to an element in $\mathbb{R}$ for $\tau=0,1,-1$, (ii) $H_{x,1}=\beta H_{x,-1}'$, and (iii) $x_{-1}^*=x_{-1}$ yields

$$
\begin{aligned}
& E_{j,-1}\left[\tilde{g}\left(x_{-1},x_0,z_0,x_1,z_1,x_2,z_2,\ldots\right)\right]-E_{j,-1}\left[\tilde{g}\left(x_{-1}^*,x_0^*,z_0,x_1^*,z_1,x_2^*,z_2,\ldots\right)\right] \\
= & \sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}\left(x_t-x_t^*\right)'H_{x,0}\left(x_t-x_t^*\right)+\left(x_t-x_t^*\right)'H_{x,1}\left(x_{t+1}-x_{t+1}^*\right)\right]. \qquad (149)
\end{aligned}
$$

Eighth, we compute the vector of first derivatives and the matrices of second derivatives appearing in equations (142) and (149). We obtain

$$
h_x'=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \qquad (150)
$$

$$
H_{x,0}=-C_j^{1-\gamma}\begin{bmatrix}
\gamma\omega_B^2\left(1+\frac{1}{\beta}\right) & \gamma\omega_B\zeta\omega_W & 0 & \cdots & 0 \\
\gamma\omega_B\zeta\omega_W & \zeta\omega_W\left(\gamma\zeta\omega_W+1+\zeta\psi\right) & 0 & \cdots & 0 \\
0 & 0 & \frac{2}{\theta I} & \cdots & \frac{1}{\theta I} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \frac{1}{\theta I} & \cdots & \frac{2}{\theta I}
\end{bmatrix}, \qquad (151)
$$

$$H_{x,1} = C_j^{1-\gamma} \begin{bmatrix} \gamma\omega_B^2 & \gamma\omega_B\zeta\omega_W & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{152}$$

$$H_{x,-1} = \frac{1}{\beta} H'_{x,1}, \tag{153}$$

$$H_{xz,0} = C_j^{1-\gamma} \left[\begin{matrix} \frac{\gamma\omega_B^2}{\beta} & -\frac{\gamma\omega_B^2}{\beta} & \frac{\gamma\omega_B\zeta^2\omega_W}{\zeta-1} & \frac{\gamma\omega_B\zeta\omega_W}{(\zeta-1)} \\ \frac{\gamma\omega_B\zeta\omega_W}{\beta} & -\frac{\gamma\omega_B\zeta\omega_W}{\beta} & \zeta^2\omega_W\left(\frac{\gamma\zeta\omega_W}{\zeta-1}+\psi\right) & \zeta\omega_W\left(\frac{\gamma\zeta\omega_W}{\zeta-1}+\psi\right) \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{matrix} \right.$$
$$\left. \begin{matrix} \gamma\omega_B\omega_D & -\gamma\omega_B\omega_T & \frac{\omega_B(1-\gamma)}{I} & \cdots & \frac{\omega_B(1-\gamma)}{I} & \frac{\omega_B(1-\gamma)}{I} \\ \gamma\zeta\omega_W\omega_D & -\gamma\zeta\omega_W\omega_T & \frac{\zeta\omega_W(1-\gamma)}{I} & \cdots & \frac{\zeta\omega_W(1-\gamma)}{I} & \frac{\zeta\omega_W(1-\gamma)}{I} \\ 0 & 0 & -\frac{1}{I} & \cdots & 0 & \frac{1}{I} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{1}{I} & \frac{1}{I} \end{matrix} \right], \tag{154}$$

$$H_{xz,1} = C_j^{1-\gamma} \left[\begin{matrix} -\frac{\gamma\omega_B^2}{\beta}+\omega_B & \frac{\gamma\omega_B^2}{\beta}-\omega_B & -\frac{\gamma\omega_B\zeta^2\omega_W}{\zeta-1} & -\frac{\gamma\omega_B\zeta\omega_W}{(\zeta-1)} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{matrix} \right.$$
$$\left. \begin{matrix} -\gamma\omega_B\omega_D & \gamma\omega_B\omega_T & -\frac{\omega_B(1-\gamma)}{I} & \cdots & -\frac{\omega_B(1-\gamma)}{I} & -\frac{\omega_B(1-\gamma)}{I} \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{matrix} \right]. \tag{155}$$

Ninth, substituting equations (150)-(155) into equation (142) yields the following system of $1+I$ equations:

$$c_{jt}^{*}=E_{t}\left[-\frac{1}{\gamma}\left(r_{t}-\pi_{t+1}-\frac{1}{I}\sum_{i=1}^{I}\left(\hat{p}_{it+1}-\hat{p}_{it}\right)\right)+c_{jt+1}^{*}\right], \tag{156}$$

$$\tilde{w}_{jt}^{*}=\frac{\gamma}{1+\zeta\psi}c_{jt}^{*}+\frac{\psi}{1+\zeta\psi}\left(\zeta\tilde{w}_{t}+l_{t}\right)+\frac{1}{1+\zeta\psi}\left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \tag{157}$$

and

$$\forall i\neq I:\hat{c}_{ijt}^{*}+\sum_{k=1}^{I-1}\hat{c}_{kjt}^{*}=-\theta\left(\hat{p}_{it}-\hat{p}_{It}\right), \tag{158}$$

where the variable c_{jt}^{*} is defined by

$$\begin{aligned} c_{jt}^{*} &= \frac{\omega_{B}}{\beta}\left(r_{t-1}-\pi_{t}+\tilde{b}_{jt-1}^{*}\right)-\omega_{B}\tilde{b}_{jt}^{*}+\frac{\zeta}{\zeta-1}\omega_{W}\left[\left(1-\zeta\right)\tilde{w}_{jt}^{*}+\zeta\tilde{w}_{t}+l_{t}\right] \\ &\quad +\omega_{D}\tilde{d}_{t}-\omega_{T}\tilde{t}_{t}-\left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right). \end{aligned} \tag{159}$$

Finally, we rewrite equation (158). Summing equation (158) over all $i\neq I$ yields

$$\sum_{i=1}^{I-1}\hat{c}_{ijt}^{*}=-\theta\left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}-\hat{p}_{It}\right).$$

Substituting the last equation back into equation (158) yields

$$\forall i\neq I:\hat{c}_{ijt}^{*}=-\theta\left(\hat{p}_{it}-\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right). \tag{160}$$

Collecting equations (149), (151), (152), (156), (157), (159) and (160), we arrive at Proposition 2.

C Solution of the model under perfect information

First, the price setting equation (47) and equations (71), (73), (76) and (12) imply that

$$0=\frac{1-\alpha}{\alpha}c_{t}+\tilde{w}_{t}-\frac{1}{\alpha}a_{t}.$$

The wage setting equation (67) and equations (71), (73) and (76) imply that

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

The production function (79) and equations (70), (72), (73) and (12) imply that

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

The equation for aggregate output (70) and equations $y_{it} = c_{it}$, (77), (68), (71) and (73) imply that

$$y_t = c_t.$$

Solving the last four equations for the endogenous variables y_t, c_t, l_t and $\tilde{w}_t$ yields equations (81)-(83). Furthermore, the consumption Euler equation (66) and equations (71) and (73) imply that

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

Substituting the solution for c_t into the last equation and solving for the real interest rate yields equation (84). Second, the equation for the optimal consumption mix (68) and equation (73) imply equation (85). Note that combining equations (85), (71) and (77) yields a demand function for good i that has the form (28)-(30) with $\vartheta = \theta$ and $\varsigma = 1$. The price setting equation (47) and equations (71), (76), (81) and (83) and a price elasticity of demand of $\vartheta = \theta$ imply equation (86). Third, the equation for the optimal factor mix (48) and equation (76) imply equation (87). Note that combining equations (87), (72) and (78) yields a labor demand function that has the form (49)-(51) with $\zeta = \eta$ and $\xi = 1$. Finally, when all households have the same initial bond holdings and the bond sequence for each household is non-explosive (i.e. $\lim_{s \to \infty} E_t \left[\beta^{s+1} \left(\tilde{b}_{j,t+s+1} - \tilde{b}_{j,t+s} \right) \right] = 0$), equations (66)-(69) have a unique solution for consumption that is identical for all households. The wage setting equation (67) then implies that all households set the same wage. It follows from equation (76) that $w_t = w_{jt}$, implying $\hat{w}_{jt} = 0$.

References

[1] Altig, David, Lawrence Christiano, Martin Eichenbaum and Jesper Linde (2004): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle." NBER Working Paper 11034.

[2] Angeletos, George-Marios and Jennifer La'O (2009): "Noisy Business Cycles." NBER Macroeconomics Annual 2009, forthcoming.

[3] Boivin, Jean, Marc Giannoni and Ilian Mihov (2009): "Sticky Prices and Monetary Policy: Evidence from Disaggregated U.S. Data." American Economic Review, 99, 350-384.

[4] Klenow, Peter J. and Oleksiy Kryvtsov (2008): "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" Quarterly Journal of Economics, 123(3), 863-904.

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

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

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

[8] Lorenzoni, Guido (2008): "A Theory of Demand Shocks." American Economic Review, forthcoming.

[9] Maćkowiak, Bartosz, Emanuel Moench and Mirko Wiederholt (2009): "Sectoral Price Data and Models of Price Setting." Journal of Monetary Economics, 56, S78-S99.

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

[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, Federal Reserve Bank of Minneapolis.

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

[14] Paciello, Luigi (2008): "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." Review of Economic Studies, forthcoming.

[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 (2009): "Information-Constrained State-Dependent Pricing." Journal of Monetary Economics, 56, S100-S124.

Figure 1: Impulse responses, benchmark economy

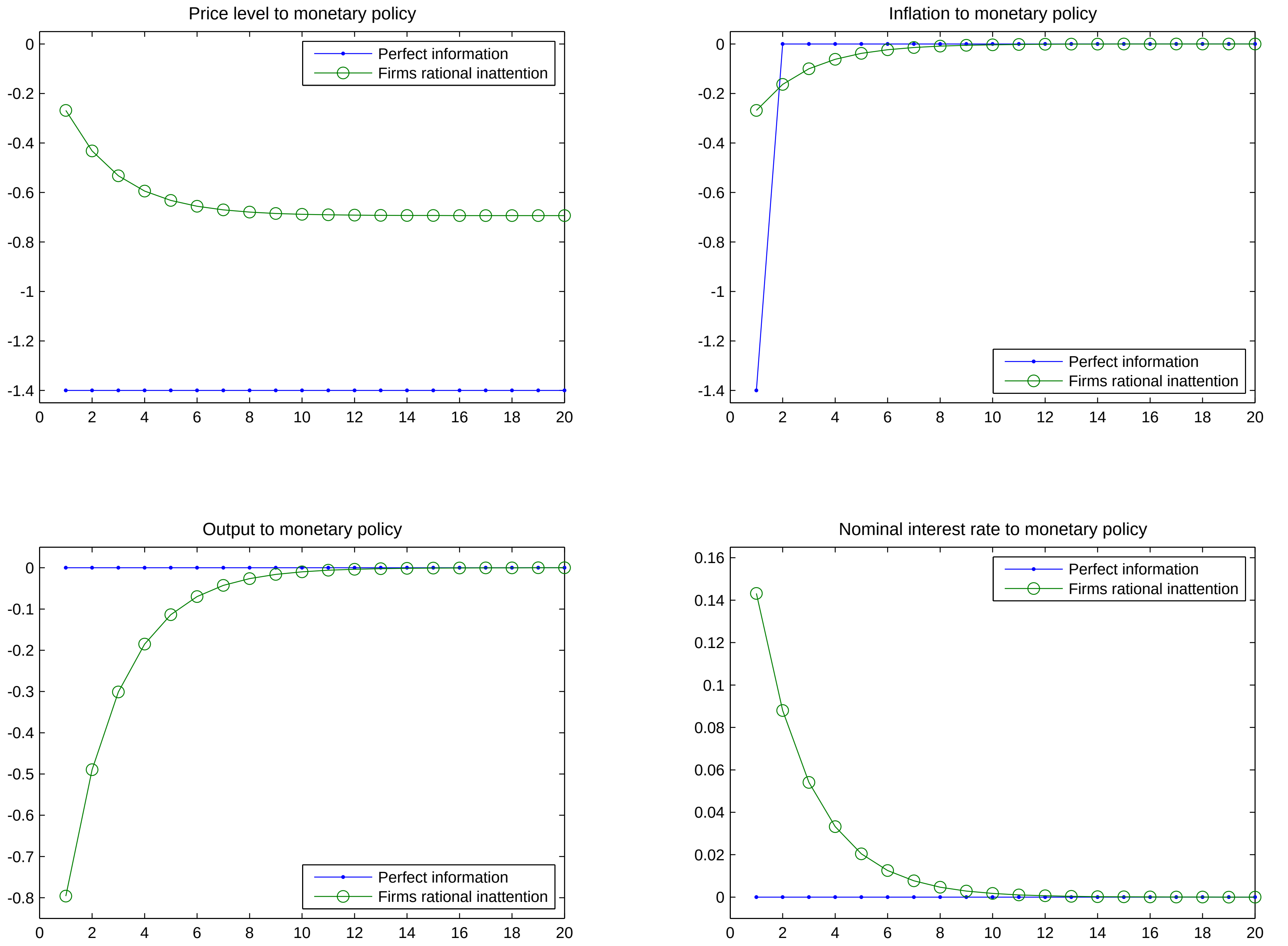


Figure 2: Impulse responses, benchmark economy

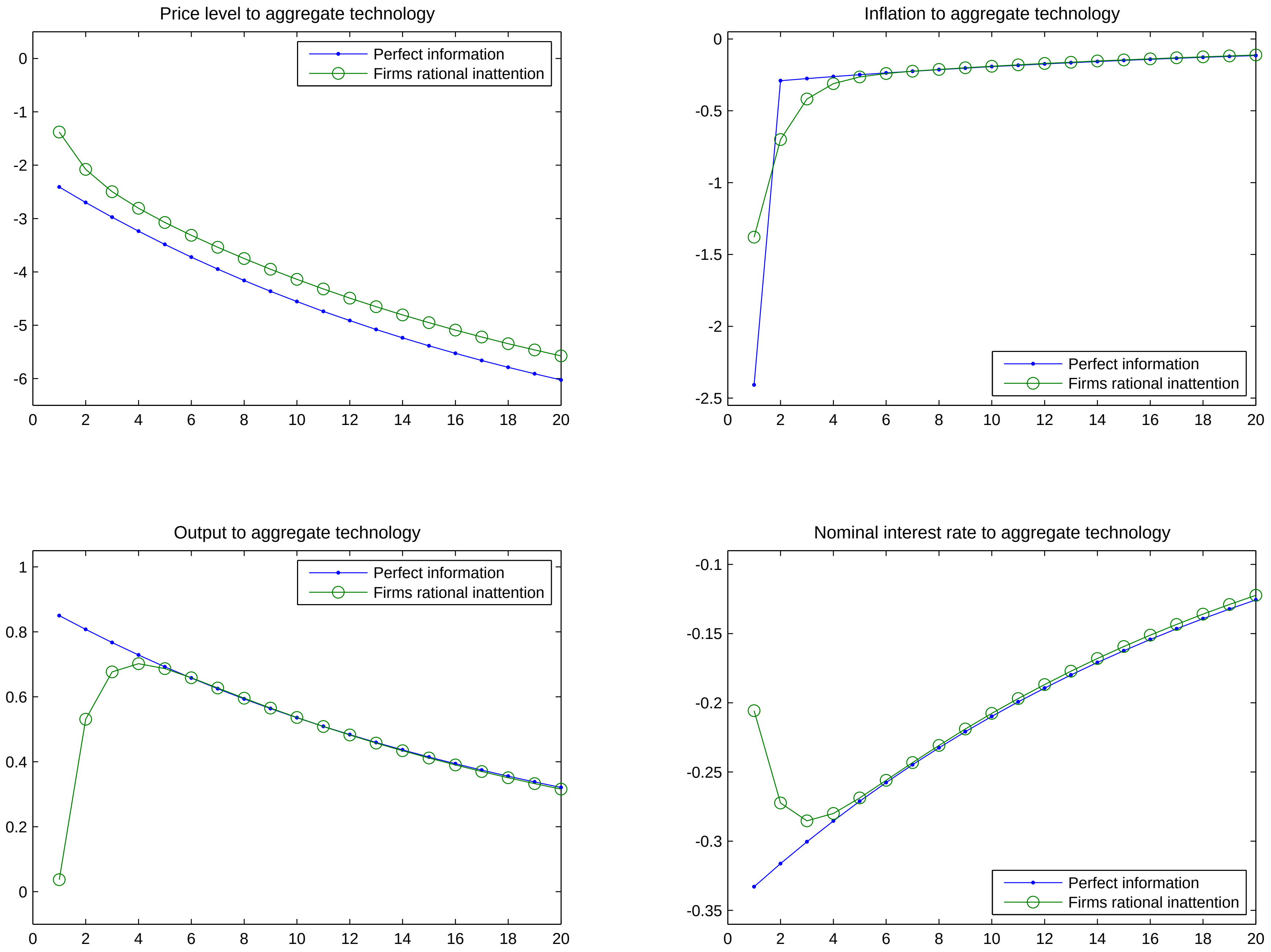
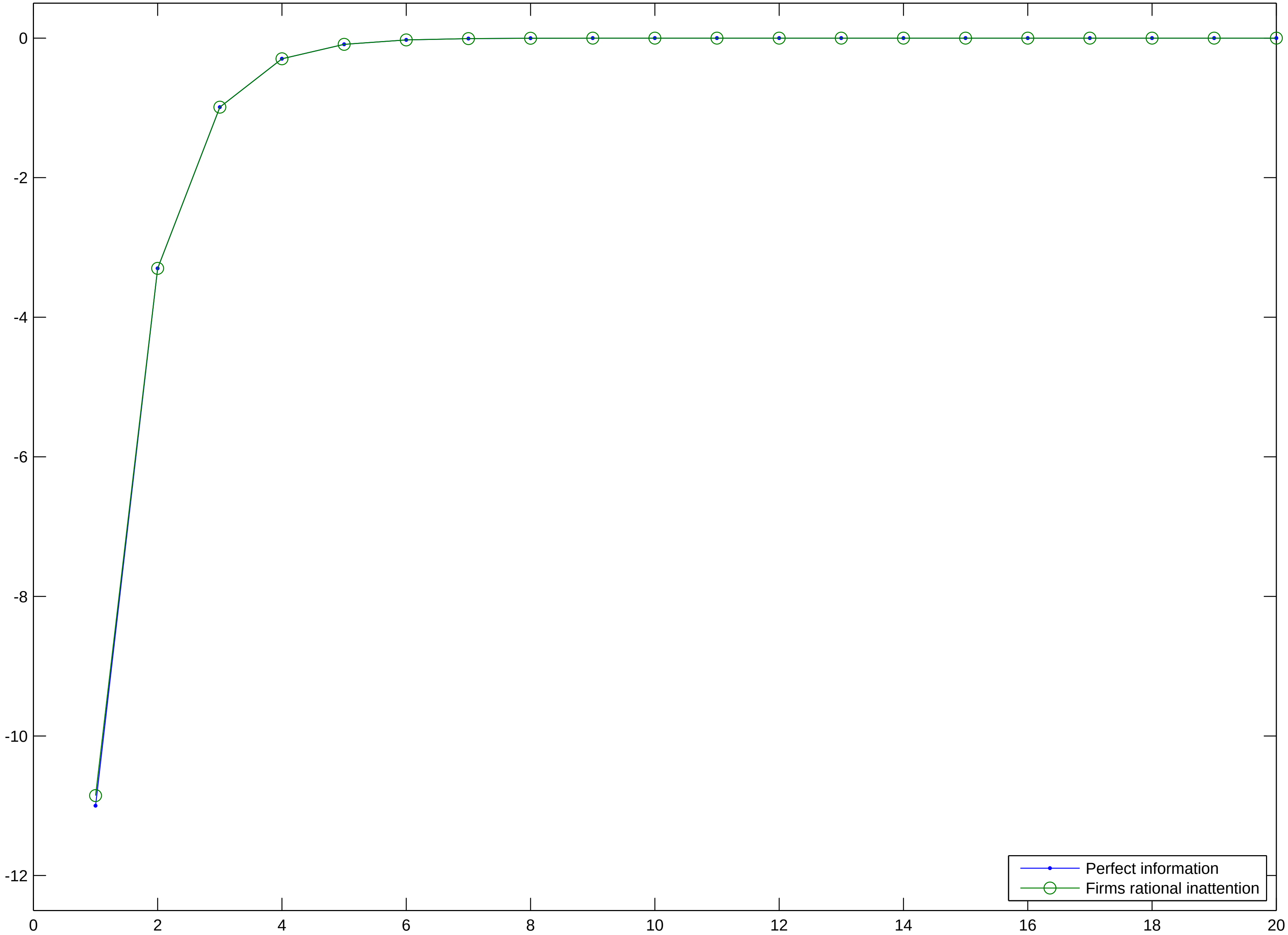


Figure 3: Impulse response of an individual price to a firm-specific productivity shock



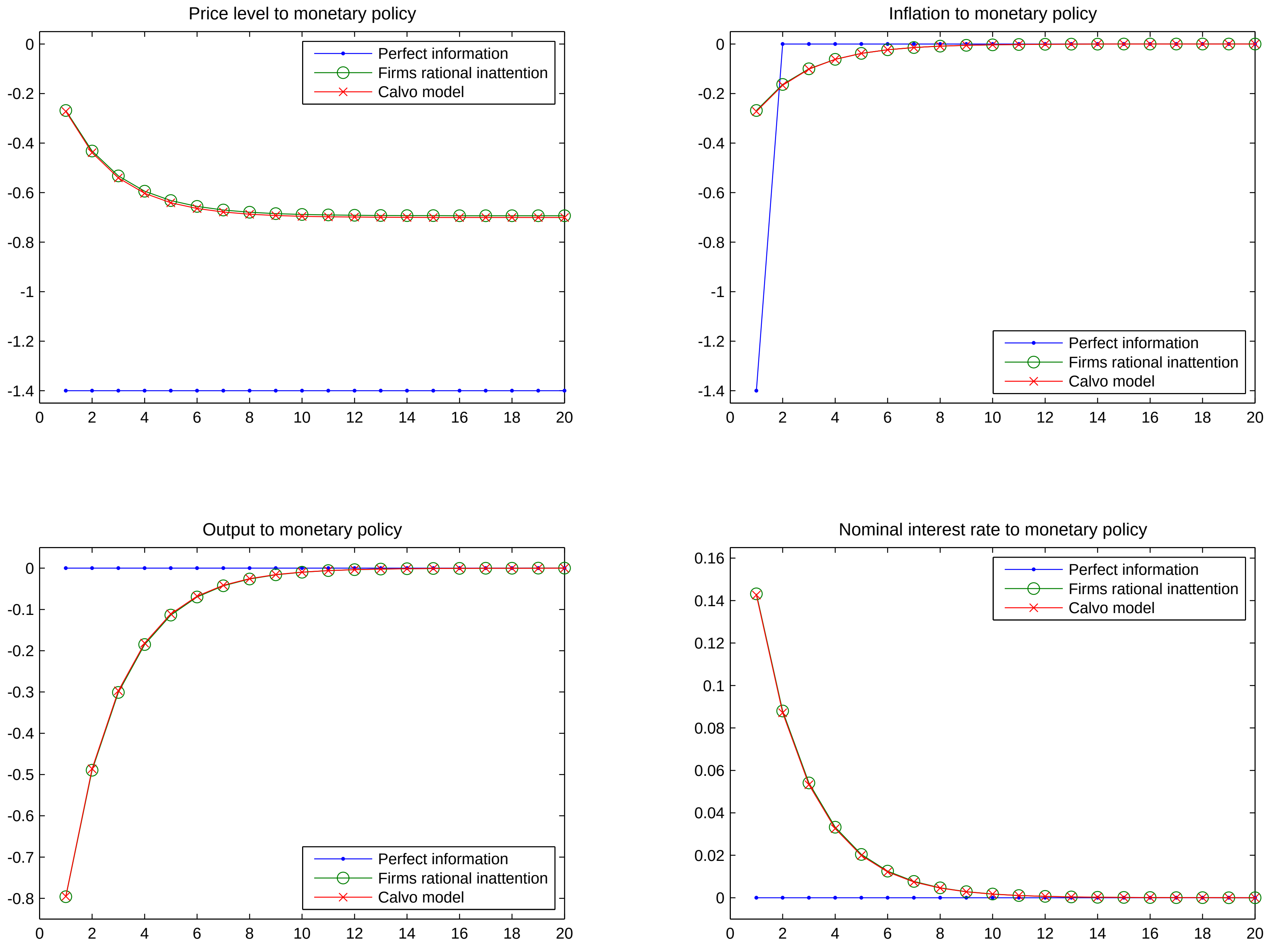


Figure 4: Impulse responses, benchmark economy and the Calvo model

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

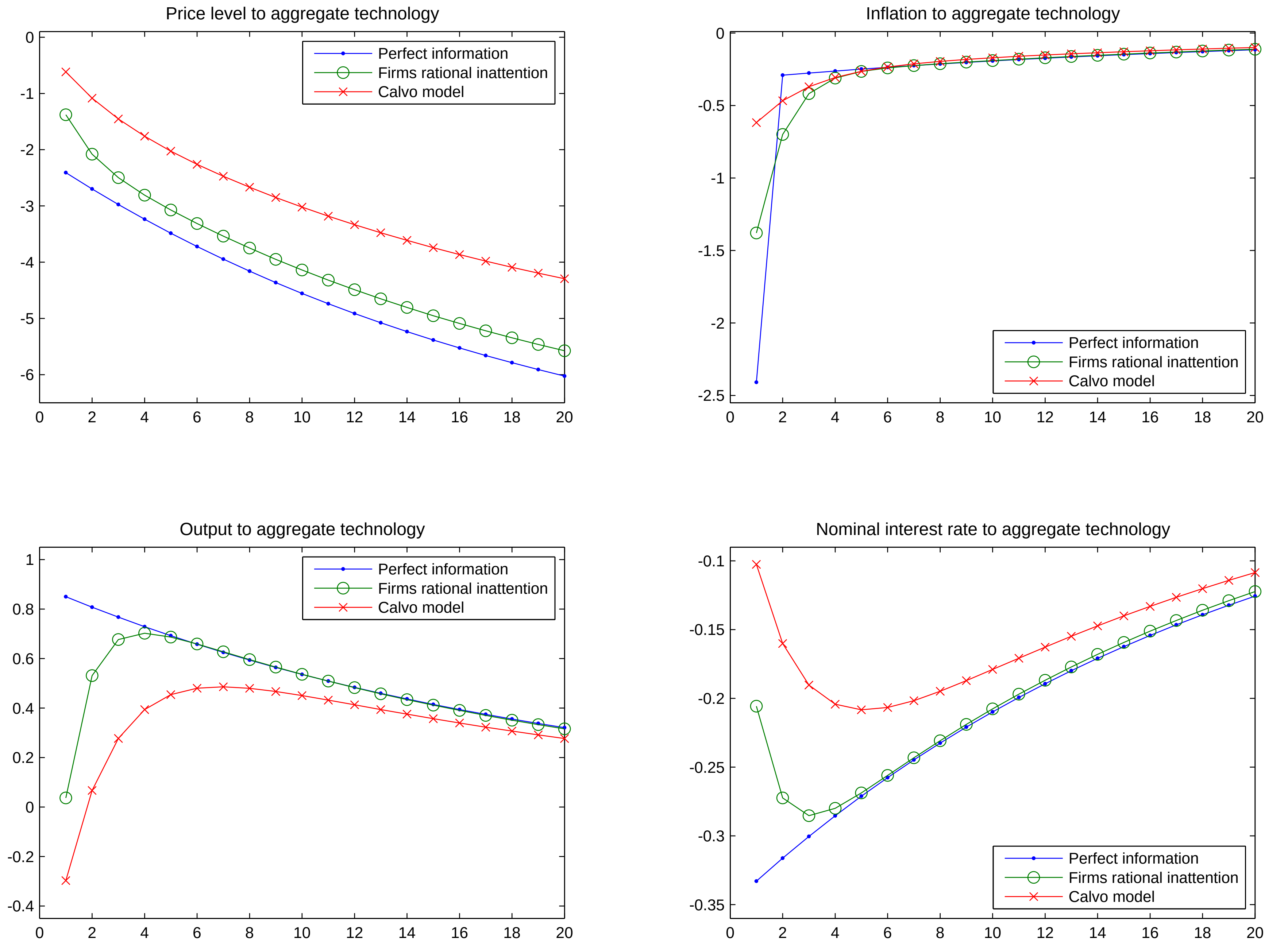


Figure 6: Impulse responses, household problem

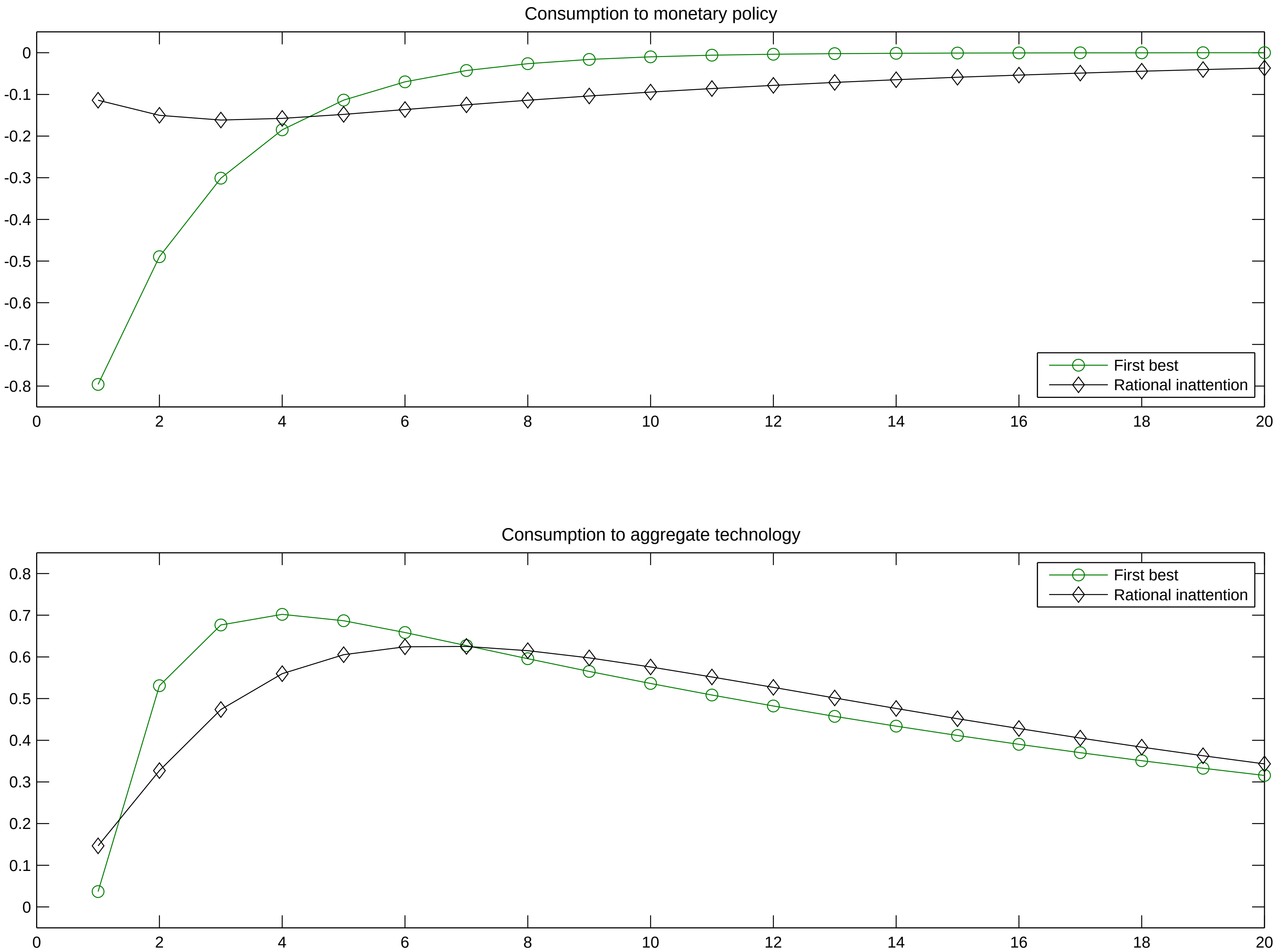


Figure 7: Impulse responses, benchmark economy

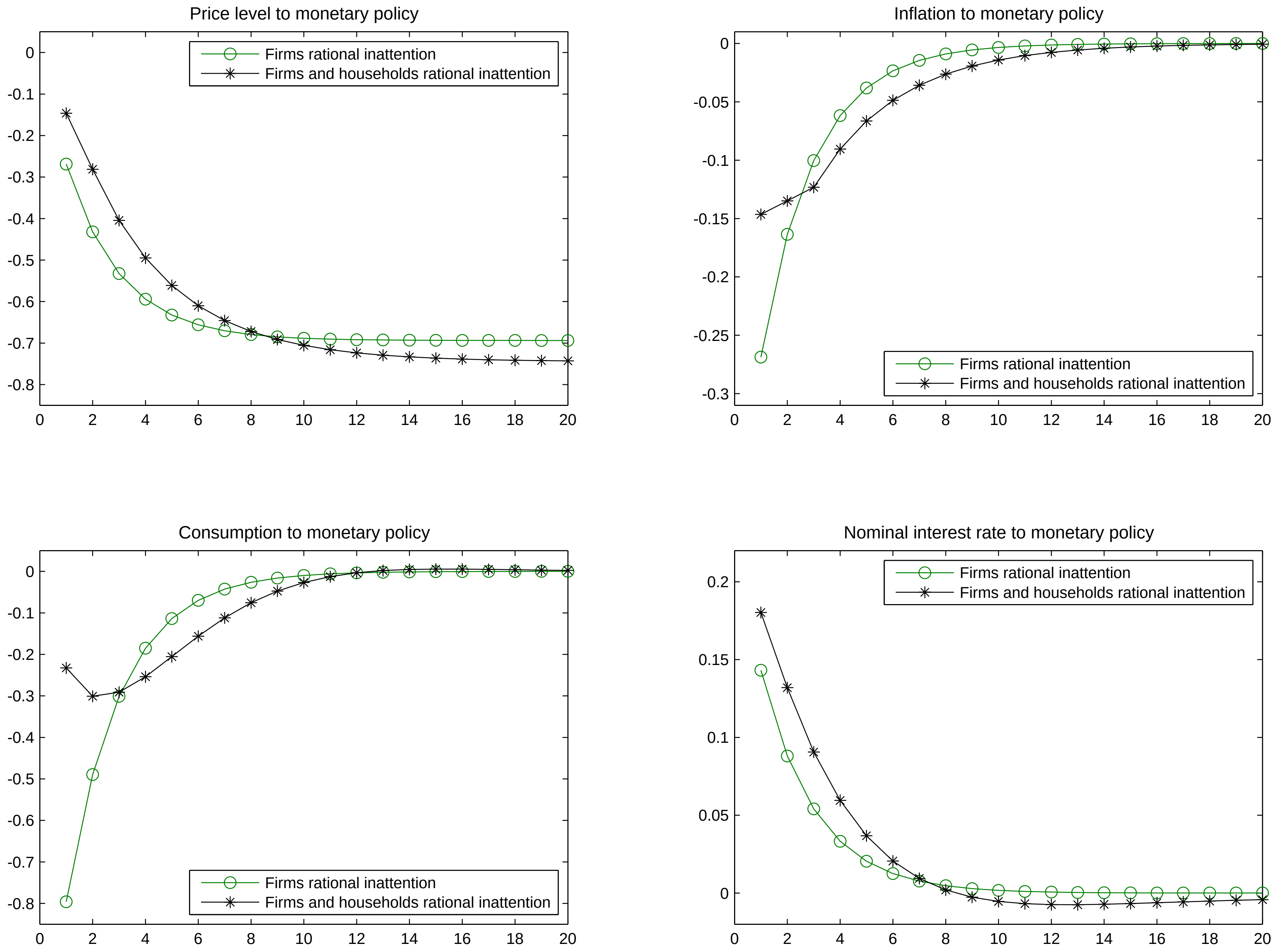


Figure 8: Impulse responses, benchmark economy