Business Cycle Dynamics under Rational Inattention*

Bartosz Maćkowiak
European Central Bank and CEPR

Mirko Wiederholt
Northwestern University

First draft: June 2007. This draft: January 2010

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. The paper studies the implications of rational inattention for business cycle dynamics. Impulse responses in the model have several properties of empirical impulse responses. Prices respond slowly to monetary policy shocks, faster to aggregate TFP shocks, and very quickly to disaggregate shocks. Therefore, 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. Consumption responds slowly to aggregate 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 and react slowly to changes in the real interest rate.

Keywords: rational inattention, information choice, dynamic stochastic general equilibrium, business cycles, monetary policy. (*JEL*: D83, E31, E32, E52).

*Maćkowiak: European Central Bank, Kaiserstrasse 29, 60311 Frankfurt am Main, Germany (e-mail: bartosz.mackowiak@ecb.int); Wiederholt: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (e-mail: m-wiederholt@northwestern.edu). We thank for helpful comments: Paco Buera, Larry Christiano, James Costain, Martin Eichenbaum, Christian Hellwig, Marek Jarociński, Giorgio Primiceri, Bruno Strulovici, Andrea Tambalotti and seminar and conference participants at Bank of Canada, Bonn, Chicago Fed, Columbia, Cowles Foundation Summer Conference 2009, DePaul, Duke, Einaudi Institute, European Central Bank, ESSIM 2008, EUI, Harvard, Madison, MIT, Minneapolis Fed, Minnesota Workshop in Macroeconomic Theory 2009, NBER Summer Institute 2008, NYU, NAWMES 2008, Philadelphia Fed, Princeton, Richmond Fed, Riksbank, SED 2008, Stony Brook, Toronto, Toulouse, UCSD, University of Chicago, University of Hong Kong, University of Montreal, Wharton and Yale. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the European Central Bank.

1 Introduction

This paper develops a dynamic stochastic general equilibrium (DSGE) model with rational inattention. Here rational inattention means that decision-makers have limited attention and decide how to allocate their attention. Following Sims (2003), we model attention as a flow of information and we model limited attention as a constraint on the flow of information. As an example, consider a household that decides how much to consume and which goods to consume. To take the optimal consumption-saving decision and to buy the optimal consumption basket, the household has to know the real interest rate and the prices of all consumption goods. The idea of rational inattention applied to this example is that: (i) knowing the real interest rate and the prices of all consumption goods requires attention, (ii) households have limited attention, and (iii) households decide how to allocate their attention. We study the implications of rational inattention for business cycle dynamics.

We are motivated by the question of how to model the inertia found in macroeconomic data. Standard DSGE models used for policy analysis match this inertia by introducing multiple sources of slow adjustment: Calvo price setting, habit formation in consumption, Calvo wage setting, and other sources in richer models.[1] We pursue the alternative idea that the inertia found in macroeconomic data can be understood as the result of rational inattention by decision-makers.

We model an economy with many firms, many households, and a government. Firms produce differentiated goods with a variety of types of labor. Households supply the differentiated types of labor, consume the different 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. The economy is affected by aggregate technology shocks, monetary policy shocks, and firm-specific productivity shocks. The only source of inertia is rational inattention by decision-makers.

We first solve the model assuming rational inattention by decision-makers in firms and perfect information on the side of households. We find that rational inattention by decision-makers in firms has the following implications: (i) the price level responds slowly to monetary policy shocks, (ii) the price level responds fairly quickly to aggregate technology shocks, and (iii) prices respond very quickly to disaggregate shocks. The reason for this combination of slow and fast adjustment of prices to shocks is that decision-makers in firms decide to pay little attention to monetary policy, more attention to aggregate technology, and a lot of attention to market-specific conditions. The

[1] See, for example, Woodford (2003), Christiano, Eichenbaum and Evans (2005), and Smets and Wouters (2007).

empirical literature finds in the data the same pattern of slow and fast adjustment of prices to shocks.[2]

In our model and in any other model with a price setting friction, firms experience profit losses due to deviations of the price from the profit-maximizing price. A nice feature of our model is that these profit losses are small. For comparison, in our benchmark economy profit losses due to deviations of the price from the profit-maximizing price are 30 times smaller than in the Calvo model that generates the same real effects of monetary policy shocks. The main reason is that in our model prices respond slowly to monetary policy shocks, fairly quickly to aggregate technology shocks, and very quickly to idiosyncratic shocks. By contrast, in the Calvo model prices respond slowly to all those shocks. The other reason is that under rational inattention deviations of the price from the profit-maximizing price are less likely to be extreme than in the Calvo model.

We use the model to conduct experiments. We find that the outcome of a policy experiment conducted in the rational inattention model can differ substantially from the outcome of the same policy experiment conducted in the Calvo model, which is the most commonly used DSGE model for policy evaluation. For example, consider increasing the coefficient on inflation in the Taylor rule. In the Calvo model, the standard deviation of the output gap due to aggregate technology shocks declines monotonically and the standard deviation of the output gap due to monetary policy shocks declines monotonically. In the rational inattention model, there is a non-monotonic relationship between output gap volatility and the coefficient on inflation in the Taylor rule. For our parameter values, the standard deviation of the output gap due to aggregate technology shocks first rises, peaking at 1.75, and then falls. The standard deviation of the output gap due to monetary policy shocks first falls, bottoming at 1.5, and then rises. The reason for the different outcomes in the two models is that in the rational inattention model there is an additional effect. When the central bank stabilizes the price level more, decision-makers in firms decide to pay less attention to aggregate conditions.

Next, we solve the model with rational inattention by decision-makers in firms and rational inattention by households. Adding rational inattention by households substantially changes the impulse responses of consumption and the nominal wage rate to aggregate shocks, despite the fact that for our parameter values a household's expected per-period loss in utility due to deviations of

[2] Christiano, Eichenbaum and Evans (1999), Leeper, Sims and Zha (1996), and Uhlig (2005) find that the price level responds slowly to monetary policy shocks. Altig, Christiano, Eichenbaum and Linde (2005) find that the price level responds faster to aggregate technology shocks than to monetary policy shocks. Boivin, Giannoni and Mihov (2009) and Maćkowiak, Moench and Wiederholt (2009) find that prices respond very quickly to disaggregate shocks.

consumption and the nominal wage rate from the optimal decisions under perfect information equals the utility equivalent of only 0.06 percent of the household's steady state consumption. Households devote little attention to aggregate conditions and respond slowly to changes in aggregate conditions because imperfect tracking of aggregate conditions causes small utility losses. Adding rational inattention by households has the following implications for the impulse responses to a monetary policy shock: (i) the impulse response of aggregate consumption becomes hump-shaped, (ii) the impulse response of the nominal wage index becomes dampened and delayed, and (iii) the impulse response of the price level becomes even more dampened and delayed, compared to the case with rational inattention by decision-makers in firms only. In addition, the impulse response of aggregate output to an aggregate technology shock becomes more dampened and delayed.

The finding that households devote little attention to aggregate conditions turns out to hold for coefficients of relative risk aversion as diverse as 1 and 10. For low values of the coefficient of relative risk aversion, deviations from the consumption Euler equation are cheap in utility terms. For high values of the coefficient of relative risk aversion, the coefficient on the real interest rate in the consumption Euler equation is small, implying that households do not want to respond strongly to changes in the real interest rate anyway. Hence, for low and high values of the coefficient of relative risk aversion, imperfect tracking of the real interest rate causes only small utility losses.

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

The paper is organized as follows. Section 2 describes all features of the economy apart from information flows. Section 3 derives the objective that decision-makers in firms maximize when they decide how to allocate their attention. Section 4 derives the objective that households maximize when they decide how to allocate their attention. Section 5 describes issues related to aggregation. Section 6 derives the solution of the model under perfect information. Section 7 presents numerical solutions of the model assuming rational inattention by decision-makers in firms and perfect information on the side of households. Section 8 presents numerical solutions of the model assuming rational inattention by decision-makers in firms and rational inattention by households.

Section 9 concludes. The appendices characterize the steady state of the non-stochastic version of the economy and contain the proofs of three propositions that appear in the main text.

2 Model

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

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.[3]

The flow budget constraint of household j in period t reads

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

Here P_{it} is the price of good i in period t, B_{jt} are 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

[3] 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 consumption aggregator.

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 will allow us to express bond holdings in terms of log-deviations from the non-stochastic steady state. One can think of households as 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, the nominal wage index defined below, the nominal interest rate and all aggregate quantities.

2.2 Firms

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

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

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

where

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

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. Type j labor is labor supplied by household j. There are J different types of labor and the parameter $\eta > 1$ is the elasticity of substitution between the types of labor. The parameter $\alpha \in (0, 1]$ is the elasticity of output with respect to composite labor input. Total factor productivity has an aggregate component, e^{a_t}, and a firm-specific component, $e^{a_{it}}$.

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 produces the quantity demanded with the chosen factor mix.

Each firm takes as given: all wage rates, the price index defined below, the nominal interest rate, all aggregate quantities and total factor productivity.

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

$$Y_t = \frac{\sum_{i=1}^{I} P_{it} Y_{it}}{P_t}, \tag{8}$$

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

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

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

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

where $\tilde{\theta}$ denotes the price elasticity of demand, and

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

where $\tilde{\eta}$ denotes the wage elasticity of labor demand.[4] We make this assumption to abstract from the level distortions arising from monopolistic competition.

[4]When households have perfect information then $\tilde{\theta} = \theta$ and $1 + \tau_p = \frac{\theta}{\theta - 1}$. By contrast, when households have imperfect information then the price elasticity of demand $\tilde{\theta}$ 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.

2.4 Shocks

There are three types of shocks in the economy: aggregate technology shocks, firm-specific productivity shocks, and monetary policy shocks. We assume that the stochastic processes $\{a_t\}$, $\{a_{1t}\}$, $\{a_{2t}\}$,..., $\{a_{It}\}$ and $\{\varepsilon_t^R\}$ are independent. Furthermore, 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.

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 Derivation of the firms' objective

In this section, we derive a log-quadratic approximation to the expected discounted sum of profits. We use this expression below when we assume that decision-makers in firms choose the allocation of their 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 good i, (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 around the non-stochastic steady state. The non-stochastic steady state of the economy presented in Section 2 is characterized in Appendix A.[5]

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

$$C_{it} = \vartheta \left(\frac{P_{it}}{P_t} \right)^{-\tilde{\theta}} C_t, \tag{19}$$

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

$$P_t = d\left(P_{1t}, \ldots, P_{It}\right), \tag{20}$$

and $\tilde{\theta} > 1$ and $\vartheta > 0$ are undetermined coefficients satisfying

$$\vartheta \hat{P}_i^{-\tilde{\theta}} = \hat{P}_i^{-\theta}. \tag{21}$$

Below when we solve the model for alternative assumptions about information flows, we always verify that this guess concerning the demand function is correct.[6]

Second, we substitute the demand function (19) and the technology (4)-(5) into the expression for profits (6) to obtain the profit function. We begin by rewriting the expression for profits (6) as follows

$$\left(1+\tau_p\right) P_{it} Y_{it} - \sum_{j=1}^{J} W_{jt} L_{ijt} = \left(1+\tau_p\right) P_{it} Y_{it} - L_{it} \left[\sum_{j=1}^{J} W_{jt} \hat{L}_{ijt} \right], \tag{22}$$

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

[5] The inflation rate in the steady state of the non-stochastic version of the economy is not uniquely determined. For ease of exposition, we select the zero inflation steady state (i.e. $\Pi = 1$). In the non-stochastic and the stochastic version of the economy, the value of Π has no effect on real variables.

[6] For example, when households have perfect information then $P_t = \left(\sum_{i=1}^{I} P_{it}^{1-\theta}\right)^{\frac{1}{1-\theta}}$, $\tilde{\theta} = \theta$ and $\vartheta = 1$.

rearranging the production function (4) and the labor aggregator (5) yields

$$L_{it} = \left(\frac{Y_{it}}{e^{a_t} e^{a_{it}}} \right)^{\frac{1}{\alpha}}, \tag{23}$$

and

$$1 = \sum_{j=1}^{J} \hat{L}_{ijt}^{\frac{\eta-1}{\eta}}. \tag{24}$$

Substituting the demand function (19), the technology (23)-(24) and $Y_{it} = C_{it}$ into the expression for profits (22) yields the profit function

$$(1+\tau_p) P_{it} \vartheta \left(\frac{P_{it}}{P_t} \right)^{-\tilde{\theta}} C_t - \left[\frac{\vartheta \left(\frac{P_{it}}{P_t} \right)^{-\tilde{\theta}} C_t}{e^{a_t} e^{a_{it}}} \right]^{\frac{1}{\alpha}} \left[\sum_{j=1}^{J-1} W_{jt} \hat{L}_{ijt} + W_{Jt} \left(1 - \sum_{j=1}^{J-1} \hat{L}_{ijt}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \right]. \tag{25}$$

Nominal profits of firm i in period t depend on variables that the decision-maker in the firm chooses: $P_{it}, \hat{L}_{i1t}, \ldots, \hat{L}_{i(J-1)t}$; and on variables that the decision-maker in the firm takes as given.

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 value profits in different states of the world. Therefore, we assume a general stochastic discount factor. We assume that, in period -1, decision-makers in firms value nominal profits in period t using the following stochastic discount factor

$$Q_{-1,t} = \beta^t \Lambda \left(C_{1t}, \ldots, C_{Jt} \right) \frac{1}{P_t}, \tag{26}$$

where P_t is the price index that appears in the demand function (19) and Λ is some twice continuously differentiable function with the property that the value of the function Λ at the non-stochastic steady state equals the marginal utility of consumption in the non-stochastic steady state[7]

$$\Lambda \left(C_1, \ldots, C_J \right) = C_j^{-\gamma}. \tag{27}$$

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

$$E_{i,-1} \left[\sum_{t=0}^{\infty} \beta^t F \left(\hat{P}_{it}, \hat{L}_{i1t}, \ldots, \hat{L}_{i(J-1)t}, a_t, a_{it}, C_{1t}, \ldots, C_{Jt}, \tilde{W}_{1t}, \ldots, \tilde{W}_{Jt} \right) \right], \tag{28}$$

[7] 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 (27) would be satisfied because all households have the same marginal utility in the non-stochastic steady state. See Appendix A.

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

$$
\begin{aligned}
& F\left(\hat{P}_{it},\hat{L}_{i1t},\ldots,\hat{L}_{i(J-1)t},a_t,a_{it},C_{1t},\ldots,C_{Jt},\tilde{W}_{1t},\ldots,\tilde{W}_{Jt}\right) \\
= \; & \Lambda\left(C_{1t},\ldots,C_{Jt}\right)\left(1+\tau_p\right)\vartheta\hat{P}_{it}^{1-\tilde{\theta}}\left(\sum_{j=1}^{J}C_{jt}\right) \\
& -\Lambda\left(C_{1t},\ldots,C_{Jt}\right)\left[\frac{\vartheta\hat{P}_{it}^{-\tilde{\theta}}\left(\sum_{j=1}^{J}C_{jt}\right)}{e^{a_t}e^{a_{it}}}\right]^{\frac{1}{\alpha}}\left[\sum_{j=1}^{J-1}\tilde{W}_{jt}\hat{L}_{ijt}+\tilde{W}_{Jt}\left(1-\sum_{j=1}^{J-1}\hat{L}_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\right]. \quad (29)
\end{aligned}
$$

We call F the real profit function. In the following, variables without the subscript t denote values in the non-stochastic steady state and small variables denote log-deviations from the non-stochastic steady state. For example, $c_{jt}=\ln\left(C_{jt}/C_j\right)$. Expressing the real profit function F in terms of log-deviations from the non-stochastic steady state and using equations (10) and (21) as well as the steady state relationships (124), (125), (127), $Y_i=L_i^{\alpha}$ and $Y_i=\hat{P}_i^{-\theta}C$ yields the following expression for the expected discounted sum of profits

$$
E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t f\left(\hat{p}_{it},\hat{l}_{i1t},\ldots,\hat{l}_{i(J-1)t},a_t,a_{it},c_{1t},\ldots,c_{Jt},\tilde{w}_{1t},\ldots,\tilde{w}_{Jt}\right)\right], \quad (30)
$$

where

$$
\begin{aligned}
& f\left(\hat{p}_{it},\hat{l}_{i1t},\ldots,\hat{l}_{i(J-1)t},a_t,a_{it},c_{1t},\ldots,c_{Jt},\tilde{w}_{1t},\ldots,\tilde{w}_{Jt}\right) \\
= \; & \Lambda\left(C_1e^{c_{1t}},\ldots,C_Je^{c_{Jt}}\right)\frac{\tilde{\theta}}{\tilde{\theta}-1}\frac{1}{\alpha}\tilde{W}L_i\frac{1}{J}\sum_{j=1}^{J}e^{\left(1-\tilde{\theta}\right)\hat{p}_{it}+c_{jt}} \\
& -\Lambda\left(C_1e^{c_{1t}},\ldots,C_Je^{c_{Jt}}\right)\tilde{W}L_ie^{-\frac{\tilde{\theta}}{\alpha}\hat{p}_{it}-\frac{1}{\alpha}\left(a_t+a_{it}\right)}\left(\frac{1}{J}\sum_{j=1}^{J}e^{c_{jt}}\right)^{\frac{1}{\alpha}} \\
& \frac{1}{J}\left[\sum_{j=1}^{J-1}e^{\tilde{w}_{jt}+\hat{l}_{ijt}}+e^{\tilde{w}_{Jt}}\left(J-\sum_{j=1}^{J-1}e^{\frac{\eta-1}{\eta}\hat{l}_{ijt}}\right)^{\frac{\eta}{\eta-1}}\right]. \quad (31)
\end{aligned}
$$

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 (31). 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*

$$\begin{aligned} x_t &= \begin{pmatrix} \hat{p}_{it} & \hat{l}_{i1t} & \cdots & \hat{l}_{i(J-1)t} \end{pmatrix}', && (32) \\ z_t &= \begin{pmatrix} a_t & a_{it} & c_{1t} & \cdots & c_{Jt} & \tilde{w}_{1t} & \cdots & \tilde{w}_{Jt} \end{pmatrix}', && (33) \\ v_t &= \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}'. && (34) \end{aligned}$$

Let $v_{m,t}$ and $v_{n,t}$ denote the mth and the nth 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,

$$E_{i,-1} \left| v_{m,t} v_{n,t} \right| < \delta^t A. \tag{35}$$

Then

$$E_{i,-1} \left[\sum_{t=0}^{\infty} \beta^t \tilde{f}(x_t, z_t) \right] - E_{i,-1} \left[\sum_{t=0}^{\infty} \beta^t \tilde{f}(x_t^*, z_t) \right] = \sum_{t=0}^{\infty} \beta^t E_{i,-1} \left[\frac{1}{2} (x_t - x_t^*)' H (x_t - x_t^*) \right], \tag{36}$$

where the matrix H is given by

$$H = -C_j^{-\gamma} \tilde{W} L_i \begin{bmatrix} \frac{\tilde{\theta}}{\alpha} \left(1 + \frac{1-\alpha}{\alpha} \tilde{\theta} \right) & 0 & \cdots & \cdots & 0 \\ 0 & \frac{2}{\eta J} & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J} \\ \vdots & \frac{1}{\eta J} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1}{\eta J} \\ 0 & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J} & \frac{2}{\eta J} \end{bmatrix}, \tag{37}$$

and the vector x_t^* is given by:

$$\hat{p}_{it}^* = \frac{\frac{1-\alpha}{\alpha}}{1 + \frac{1-\alpha}{\alpha} \tilde{\theta}} \left(\frac{1}{J} \sum_{j=1}^{J} c_{jt} \right) + \frac{1}{1 + \frac{1-\alpha}{\alpha} \tilde{\theta}} \left(\frac{1}{J} \sum_{j=1}^{J} \tilde{w}_{jt} \right) - \frac{\frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha} \tilde{\theta}} (a_t + a_{it}), \tag{38}$$

and

$$\hat{l}_{ijt}^* = -\eta \left(\tilde{w}_{jt} - \frac{1}{J} \sum_{j=1}^{J} \tilde{w}_{jt} \right). \tag{39}$$

Proof. See Appendix B. ■

After the log-quadratic approximation to the real profit function, the profit-maximizing price in period t is given by equation (38) and the profit-maximizing factor mix in period t is given by equation (39). In addition, the loss in profits in period t in the case of a deviation from the profit-maximizing decision vector is given by the quadratic form in expression (36). The upper-left element of the matrix H determines the loss in profits in the case of a suboptimal price. The profit loss in the case of a suboptimal price is increasing in the price elasticity of demand, $\tilde{\theta}$, and increasing in the degree of decreasing returns-to-scale, $(1/\alpha)$. The lower-right block of the matrix H determines the loss in profits in the case of a suboptimal factor mix. The profit loss in the case of a suboptimal factor mix 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 (35) 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.

Note that the profit-maximizing decision vector (38)-(39) does not depend at all on the function Λ appearing in the stochastic discount factor (26) because the profit-maximizing price and the profit-maximizing factor mix are the solution to a static maximization problem. Furthermore, the expected discounted sum of profit losses (36) depends only on the value of the function Λ in the non-stochastic steady state because of the log-quadratic approximation to the real profit function around the non-stochastic steady state.

Proposition 1 gives an expression for the expected discounted sum of profits for the economy presented in Section 2 when the demand function is given by equation (19) and the stochastic discount factor is given by equation (26). 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 or 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 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 (38)-(39), 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.

4 Derivation of the households' objective

Next, we derive a log-quadratic approximation to the expected discounted sum of period utility. We use this expression below when we assume that households choose the allocation of their 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 type j 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 around the non-stochastic steady state.

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

$$L_{jt} = \zeta \left(\frac{W_{jt}}{W_t} \right)^{-\tilde{\eta}} L_t, \tag{40}$$

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

$$W_t = h\left(W_{1t}, \ldots, W_{Jt}\right), \tag{41}$$

and $\tilde{\eta} > 1$ and $\zeta > 0$ are undetermined coefficients satisfying

$$\zeta \hat{W}_j^{-\tilde{\eta}} = \hat{W}_j^{-\eta}. \tag{42}$$

Below when we solve the model for alternative assumptions about information flows, we always verify that this guess concerning the labor demand function is correct.[8]

Second, we substitute the consumption aggregator (2), the flow budget constraint (3) and the labor demand function (40) 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

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

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

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

[8] For example, when firms have perfect information then $W_t = \left(\sum_{j=1}^{J} W_{jt}^{1-\eta} \right)^{\frac{1}{1-\eta}}$, $\tilde{\eta} = \eta$ and $\zeta = 1$.

Dividing the numerator and the denominator on the right-hand side of the last equation by some price index P_t yields

$$C_{jt} = \frac{\frac{R_{t-1}}{\Pi_t}\tilde{B}_{jt-1} - \tilde{B}_{jt} + (1+\tau_w)\tilde{W}_{jt}L_{jt} + \frac{\tilde{D}_t}{J} - \frac{\tilde{T}_t}{J}}{\sum_{i=1}^{I}\hat{P}_{it}\hat{C}_{ijt}}, \tag{43}$$

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. Furthermore, rearranging the consumption aggregator (2) yields

$$1 = \sum_{i=1}^{I}\hat{C}_{ijt}^{\frac{\theta-1}{\theta}}. \tag{44}$$

Substituting the labor demand function (40), the flow budget constraint (43) and the consumption aggregator (44) into the period utility function (1) yields a period utility function that incorporates those constraints:

$$\frac{1}{1-\gamma}\left(\frac{\frac{R_{t-1}}{\Pi_t}\tilde{B}_{jt-1} - \tilde{B}_{jt} + (1+\tau_w)\tilde{W}_{jt}\zeta\left(\frac{\tilde{W}_{jt}}{\tilde{W}_t}\right)^{-\tilde{\eta}}L_t + \frac{\tilde{D}_t}{J} - \frac{\tilde{T}_t}{J}}{\sum_{i=1}^{I-1}\hat{P}_{it}\hat{C}_{ijt} + \hat{P}_{It}\left(1 - \sum_{i=1}^{I-1}\hat{C}_{ijt}^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}}\right)^{1-\gamma}$$
$$-\frac{1}{1-\gamma} - \frac{\varphi}{1+\psi}\left[\zeta\left(\frac{\tilde{W}_{jt}}{\tilde{W}_t}\right)^{-\tilde{\eta}}L_t\right]^{1+\psi}. \tag{45}$$

Expressing the period utility function (45) in terms of log-deviations from the non-stochastic steady state and using equations (11) and (42) as well as the steady state relationships (121)-(123), (126) and $L_j = \hat{W}_j^{-\eta}L$ yields the following period utility function

$$\frac{C_j^{1-\gamma}}{1-\gamma}\left(\frac{\frac{\omega_B}{\beta}e^{r_{t-1}-\pi_t+\tilde{b}_{jt-1}} - \omega_B e^{\tilde{b}_{jt}} + \frac{\tilde{\eta}}{\tilde{\eta}-1}\omega_W e^{(1-\tilde{\eta})\tilde{w}_{jt}+\tilde{\eta}\tilde{w}_t+l_t} + \omega_D e^{\tilde{d}_t} - \omega_T e^{\tilde{t}_t}}{\frac{1}{I}\sum_{i=1}^{I-1}e^{\hat{p}_{it}+\hat{c}_{ijt}} + \frac{1}{I}e^{\hat{p}_{It}}\left(I - \sum_{i=1}^{I-1}e^{\frac{\theta-1}{\theta}\hat{c}_{ijt}}\right)^{\frac{\theta}{\theta-1}}}\right)^{1-\gamma}$$
$$-\frac{1}{1-\gamma} - \frac{C_j^{1-\gamma}}{1+\psi}\omega_W e^{-\tilde{\eta}(1+\psi)(\tilde{w}_{jt}-\tilde{w}_t)+(1+\psi)l_t}, \tag{46}$$

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

$$\begin{pmatrix}\omega_B & \omega_W & \omega_D & \omega_T\end{pmatrix} = \begin{pmatrix}\frac{\tilde{B}_j}{C_j} & \frac{\tilde{W}_j L_j}{C_j} & \frac{\frac{\tilde{D}}{J}}{C_j} & \frac{\frac{\tilde{T}}{J}}{C_j}\end{pmatrix}. \tag{47}$$

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 (46) 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*

$$x_t = \begin{pmatrix} \tilde{b}_{jt} & \tilde{w}_{jt} & \hat{c}_{1jt} & \cdots & \hat{c}_{I-1jt} \end{pmatrix}', \tag{48}$$

$$z_t = \begin{pmatrix} r_{t-1} & \pi_t & \tilde{w}_t & l_t & \tilde{d}_t & \tilde{t}_t & \hat{p}_{1t} & \cdots & \hat{p}_{It} \end{pmatrix}', \tag{49}$$

$$v_t = \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}'. \tag{50}$$

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

$$E_{j,-1}\left[\tilde{b}_{j,-1}^2\right] < \infty, \tag{51}$$

and, for all m,

$$E_{j,-1}\left|\tilde{b}_{j,-1} v_{m,0}\right| < \infty. \tag{52}$$

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,

$$E_{j,-1}\left|v_{m,t} v_{n,t+\tau}\right| < \delta^t A. \tag{53}$$

Then

$$\begin{aligned} & E_{j,-1}\left[\tilde{g}\left(\tilde{b}_{j,-1}, x_0, z_0, x_1, z_1, \ldots\right)\right] - E_{j,-1}\left[\tilde{g}\left(\tilde{b}_{j,-1}, x_0^*, z_0, x_1^*, z_1, \ldots\right)\right] \\ = & \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2}\left(x_t - x_t^*\right)' H_0 \left(x_t - x_t^*\right) + \left(x_t - x_t^*\right)' H_1 \left(x_{t+1} - x_{t+1}^*\right)\right]. \end{aligned} \tag{54}$$

Here the matrix H_0 is given by

$$H_0 = -C_j^{1-\gamma} \begin{bmatrix} \gamma\omega_B^2\left(1+\frac{1}{\beta}\right) & \gamma\omega_B\tilde{\eta}\omega_W & 0 & \cdots & 0 \\ \gamma\omega_B\tilde{\eta}\omega_W & \tilde{\eta}\omega_W\left(\gamma\tilde{\eta}\omega_W + 1 + \tilde{\eta}\psi\right) & 0 & \cdots & 0 \\ 0 & 0 & \frac{2}{\theta I} & \cdots & \frac{1}{\theta I} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \frac{1}{\theta I} & \cdots & \frac{2}{\theta I} \end{bmatrix}, \tag{55}$$

the matrix H_1 is given by

$$H_1 = C_j^{1-\gamma} \begin{bmatrix} \gamma\omega_B^2 & \gamma\omega_B\tilde{\eta}\omega_W & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{56}$$

and the stochastic process $\{x_t^*\}_{t=0}^{\infty}$ is defined by the following three requirements: (i) $\tilde{b}_{j,-1}^* = \tilde{b}_{j,-1}$, (ii) in every period $t \geq 0$, the vector x_t^* satisfies

$$c_{jt}^* = E_t\left[-\frac{1}{\gamma}\left(r_t - \pi_{t+1} - \frac{1}{I}\sum_{i=1}^{I}\left(\hat{p}_{it+1} - \hat{p}_{it}\right)\right) + c_{jt+1}^*\right], \tag{57}$$

$$\tilde{w}_{jt}^* = \frac{\gamma}{1+\tilde{\eta}\psi}c_{jt}^* + \frac{\psi}{1+\tilde{\eta}\psi}\left(\tilde{\eta}\tilde{w}_t + l_t\right) + \frac{1}{1+\tilde{\eta}\psi}\left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \tag{58}$$

$$\hat{c}_{ijt}^* = -\theta\left(\hat{p}_{it} - \frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \tag{59}$$

where E_t denotes the expectation operator conditioned on the entire history of the economy up to and including period t and the variable c_{jt}^* is defined by

$$c_{jt}^* = \frac{\omega_B}{\beta}\left(r_{t-1} - \pi_t + \tilde{b}_{jt-1}^*\right) - \omega_B\tilde{b}_{jt}^* + \frac{\tilde{\eta}}{\tilde{\eta}-1}\omega_W\left[\left(1-\tilde{\eta}\right)\tilde{w}_{jt}^* + \tilde{\eta}\tilde{w}_t + l_t\right] + \omega_D\tilde{d}_t - \omega_T\tilde{t}_t - \left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \tag{60}$$

and (iii) the vector v_t with $x_t = x_t^*$ satisfies conditions (51)-(53).

Proof. See Appendix C. ■

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 satisfying conditions (51)-(53) can be ranked using equation (54). Equations (57)-(60) characterize the optimal decisions under perfect information. More precisely, equations (57)-(60) characterize the decisions that the same household with the same initial bond holdings would take if the household had perfect information in every period $t \geq 0$. Equation (54) gives the loss in expected utility due to deviations from the optimal decisions 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 utility due to suboptimal real bond holdings and real wage rates. A single deviation of real bond holdings from optimal real 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 deviation of the real wage rate from the optimal real wage rate causes a larger utility loss the larger γ, ψ, ω_W and $\tilde{\eta}$. 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 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 loss in utility in the case of a suboptimal consumption mix. The loss in utility in the case of a suboptimal consumption mix is decreasing in the elasticity of substitution between consumption goods, θ, and depends on the number of consumption goods, I. Finally, conditions (51)-(53) ensure that, in the expression for the expected discounted sum of period utility, after the log-quadratic approximation, one can change the order of integration and summation and all infinite sums converge.

Proposition 2 gives an expression for the expected discounted sum of period utility for the economy presented in Section 2 when the labor demand function is given by equation (40). From this expression one can already see how some parameters will affect the optimal allocation of attention by a household that cannot attend perfectly to all available information. For example, consider the role of γ. Raising γ increases the utility loss in the case of a given deviation of real bond holdings from optimal real bond holdings. On the other hand, raising γ lowers the response of optimal real 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.

5 Aggregation

In this section, we describe issues related to aggregation. In the following, we 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

$$y_t = \frac{1}{I}\sum_{i=1}^{I}\left(\hat{p}_{it} + y_{it}\right), \tag{61}$$

$$c_t = \frac{1}{J}\sum_{j=1}^{J} c_{jt}, \tag{62}$$

and

$$l_t = \frac{1}{I}\sum_{i=1}^{I} l_{it}. \tag{63}$$

Log-linearizing the equations for the price index (20) and for the wage index (41) yields

$$0 = \sum_{i=1}^{I} \hat{p}_{it}, \tag{64}$$

and

$$0 = \sum_{j=1}^{J} \hat{w}_{jt}. \tag{65}$$

Note that the last two equations can also be stated as

$$p_t = \frac{1}{I} \sum_{i=1}^{I} p_{it}, \tag{66}$$

and

$$w_t = \frac{1}{J} \sum_{j=1}^{J} w_{jt}. \tag{67}$$

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

$$c_{it} = \frac{1}{J} \sum_{j=1}^{J} c_{ijt}, \tag{68}$$

and

$$l_{jt} = \frac{1}{I} \sum_{i=1}^{I} l_{ijt}. \tag{69}$$

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

$$y_{it} = a_t + a_{it} + \alpha l_{it}, \tag{70}$$

and

$$r_t = \rho_R r_{t-1} + (1 - \rho_R) \left(\phi_\pi \pi_t + \phi_y y_t \right) + \varepsilon_t^R. \tag{71}$$

6 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 3), the log-quadratic approximation to the expected discounted sum of period utility (see Section 4), and the log-linearization of the equations for the aggregate variables (see Section 5).

Proposition 3 *(Solution of the model under perfect information) A solution to the system of equations (38)-(39), (57)-(60), (61)-(71), (12) and* $y_{it} = c_{it}$ *with the same initial bond holdings 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:*

$$\begin{aligned} y_t &= c_t = \frac{1+\psi}{1-\alpha+\alpha\gamma+\psi} a_t, && (72)\\ l_t &= \frac{1-\gamma}{1-\alpha+\alpha\gamma+\psi} a_t, && (73)\\ \tilde{w}_t &= \frac{\gamma+\psi}{1-\alpha+\alpha\gamma+\psi} a_t, && (74)\\ r_t - E_t\left[\pi_{t+1}\right] &= \gamma\frac{1+\psi}{1-\alpha+\alpha\gamma+\psi} E_t\left[a_{t+1} - a_t\right], && (75) \end{aligned}$$

and

$$\begin{aligned} \hat{c}_{ijt} &= -\theta\hat{p}_{it}, && (76)\\ \hat{p}_{it} &= -\frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\theta} a_{it}, && (77)\\ \hat{l}_{ijt} &= -\eta\hat{w}_{jt}, && (78)\\ \hat{w}_{jt} &= 0. && (79) \end{aligned}$$

Proof. See Appendix D. ■

Under perfect information, aggregate output, aggregate consumption, aggregate labor input, the real wage index and the real interest rate depend only on aggregate technology. Furthermore, relative consumption of good i by household j depends only on firm-specific productivity of firm i. In addition, firm i's relative input of type j labor is constant. Under perfect information, monetary policy has no effect on real variables in this model. Monetary policy does affect nominal variables. The nominal interest rate and inflation follow from the monetary policy rule (71) and the real interest rate (75). 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.[9]

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

7 Rational inattention by firms

In this section, we solve the model assuming rational inattention by decision-makers in firms. For the moment, we maintain the assumption that households have perfect information to isolate the implications of rational inattention by decision-makers in firms.

7.1 The firms' attention problem

Following Sims (2003), we model attention as a flow of information and we model limited attention as a constraint on the flow of information. We let decision-makers choose information flows, subject to the constraint on information flow.

To take decisions that are close to the profit-maximizing decisions, decision-makers in firms have to be aware of changes in the environment that cause changes in the profit-maximizing decisions. Being aware of stochastic changes in the environment requires information flow. A decision-maker with limited attention faces a trade-off: Tracking closely particular changes in the environment improves decision making but 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 in a firm can decide to respond swiftly and correctly with the price of the good to changes in firm-specific productivity but this requires 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:

$$\max_{\kappa,B_1(L),B_2(L),B_3(L),C_1(L),C_2(L),C_3(L),\tilde{\eta},\chi} \left\{ \sum_{t=0}^{\infty} \beta^t E_{i,-1} \left[\frac{1}{2} \left(x_t - x_t^* \right)' H \left(x_t - x_t^* \right) \right] - \frac{\mu}{1-\beta} \kappa \right\}, \tag{80}$$

where

$$x_t - x_t^* = \begin{pmatrix} p_{it} \\ \hat{l}_{i1t} \\ \vdots \\ \hat{l}_{i(J-1)t} \end{pmatrix} - \begin{pmatrix} p_{it}^* \\ \hat{l}_{i1t}^* \\ \vdots \\ \hat{l}_{i(J-1)t}^* \end{pmatrix}, \tag{81}$$

subject to the equations characterizing the profit-maximizing decisions

$$p_{it}^* = \underbrace{A_1(L)\,\varepsilon_t^A}_{p_{it}^{A*}} + \underbrace{A_2(L)\,\varepsilon_t^R}_{p_{it}^{R*}} + \underbrace{A_3(L)\,\varepsilon_{it}^I}_{p_{it}^{I*}} \tag{82}$$

$$\hat{l}_{ijt}^* = -\eta \hat{w}_{jt}, \tag{83}$$

the equations characterizing the actual decisions

$$p_{it} = \underbrace{B_1(L)\,\varepsilon_t^A + C_1(L)\,\nu_{it}^A}_{p_{it}^A} + \underbrace{B_2(L)\,\varepsilon_t^R + C_2(L)\,\nu_{it}^R}_{p_{it}^R} + \underbrace{B_3(L)\,\varepsilon_{it}^I + C_3(L)\,\nu_{it}^I}_{p_{it}^I} \tag{84}$$

$$\hat{l}_{ijt} = -\tilde{\eta}\left(\hat{w}_{jt} + \frac{Var(\hat{w}_{jt})}{\chi}\nu_{ijt}^L\right), \tag{85}$$

and the constraint on information flow

$$\mathcal{I}\left(\left\{p_{it}^{A*}, p_{it}^{R*}, p_{it}^{I*}, \hat{l}_{i1t}^{*}, \ldots, \hat{l}_{i(J-1)t}^{*}\right\}; \left\{p_{it}^{A}, p_{it}^{R}, p_{it}^{I}, \hat{l}_{i1t}, \ldots, \hat{l}_{i(J-1)t}\right\}\right) \leq \kappa. \tag{86}$$

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 defined below. Finally, $E_{i,-1}$ in objective (80) denotes the expectation operator conditioned on the information of the decision-maker of firm i in period -1. We assume that $E_{i,-1}$ is the unconditional expectation operator.

The objective (80) states that the decision-maker in 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. Recall that after the log-quadratic approximation to the real profit function, the expected discounted sum of losses in profits due to suboptimal decisions is given by equation (36). See Proposition 1.[10] The parameter $\mu \geq 0$ is the per-period marginal cost of information flow. The variable $\kappa \geq 0$ is the overall information flow devoted to the price setting decision and the factor mix decision. The marginal cost of information flow μ can be interpreted as an opportunity cost (i.e. devoting more attention to the price setting decision or the factor mix decision requires paying less attention to some other decision) or a monetary cost (e.g. a wage payment).

Equations (82)-(83) characterize the profit-maximizing decisions. After the log-quadratic approximation to the real profit function, the profit-maximizing price is given by equation (38) and the profit-maximizing factor mix is given by equation (39). Here we guess that the profit-maximizing price (38) has the representation (82) after using equations (62), (67) and $\hat{p}_{it} = p_{it} - p_t$ and after substituting in the equilibrium law of motion for p_t, c_t, $\tilde{w}_t$, a_t and a_{it}. The guess will be verified. Furthermore, rewriting the equation for the profit-maximizing factor mix (39) using equations (67) and $\hat{w}_{jt} = \tilde{w}_{jt} - \tilde{w}_t$ yields equation (83).

[10] In equation (81), we use the fact that $\hat{p}_{it} - \hat{p}_{it}^* = p_{it} - p_{it}^*$.

Equations (84)-(85) characterize the actual decisions. Consider first equation (84). 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 stochastic process for the price. For example, if the decision-maker chooses $B_1(L) = A_1(L)$, $C_1(L) = 0$, $B_2(L) = A_2(L)$, $C_2(L) = 0$, $B_3(L) = A_3(L)$ and $C_3(L) = 0$, the decision-maker decides to set the profit-maximizing price in each period. The basic trade-off is the following. Choosing a process for the price that tracks more closely the profit-maximizing price reduces the expected loss in profits due to deviations of the price from the profit-maximizing price but requires a larger information flow. Consider next equation (85). By choosing the coefficients $\tilde{\eta}$ and χ, the decision-maker chooses the wage elasticity of labor demand and the signal-to-noise ratio in the factor mix decision. The basic trade-off is the following. Choosing a process for the factor mix that tracks more closely the profit-maximizing factor mix reduces the expected loss in profits due to deviations of the factor mix from the profit-maximizing factor mix but requires a larger information flow so long as the profit-maximizing factor mix is stochastic.[11]

The constraint (86) states that actual decisions containing more information about the profit-maximizing decisions (i.e. the optimal decisions under perfect information) require a larger information flow.

We follow Sims (2003) and a large literature in information theory by quantifying information flow as reduction in uncertainty, where uncertainty is measured by entropy. Formally, let $H(X)$ denote the entropy of the random vector $X = (X_1, \ldots, X_N)$. Entropy is a measure of uncertainty. Furthermore, let $H(X|Y)$ denote the conditional entropy of the random vector X given knowledge of $Y = (Y_1, \ldots, Y_N)$. Conditional entropy is a measure of conditional uncertainty. Equipped with measures of uncertainty and conditional uncertainty, one can quantify the information that Y contains about X by the reduction in uncertainty $H(X) - H(X|Y)$. More precisely, the operator $\mathcal{I}$ in the information flow constraint (86) is defined as

$$\mathcal{I}(\{X_t\};\{Y_t\}) = \lim_{T\to\infty}\frac{1}{T}\left[H(X_0,\ldots,X_{T-1}) - H(X_0,\ldots,X_{T-1}|Y_0,\ldots,Y_{T-1})\right], \tag{87}$$

where $\{X_t\}_{t=0}^{\infty}$ and $\{Y_t\}_{t=0}^{\infty}$ are stochastic processes. Thus, we quantify the information that one process, $\{Y_t\}_{t=0}^{\infty}$, contains about another process, $\{X_t\}_{t=0}^{\infty}$, by measuring the average per-period amount of information that the first T elements of one process contain about the first T elements

[11] We put more structure on the factor mix decision than on the price setting decision. In particular, in equation (85) we express the factor mix as a function of relative wage rates rather than expressing the factor mix as a function of fundamental shocks. We do this because from equation (85) we derive the labor demand function and a labor demand function specifies labor demand on and off the equilibrium path. By expressing the labor mix as a function of relative wage rates we specify firm i's relative input of type j labor on and off the equilibrium path.

of the other process and by letting T go to infinity. If $\{X_t, Y_t\}_{t=0}^{\infty}$ is a stationary Gaussian process, then the difference between entropy and conditional entropy on the right-hand side of equation (87) simply equals

$$H\left(X_0,\ldots,X_{T-1}\right)-H\left(X_0,\ldots,X_{T-1}|Y_0,\ldots,Y_{T-1}\right)=\frac{1}{2}\log_2\left(\frac{\det\Omega_X}{\det\Omega_{X|Y}}\right). \tag{88}$$

Here $\det\Omega_X$ denotes the determinant of the covariance matrix of $(X_0,\ldots,X_{T-1})$ and $\det\Omega_{X|Y}$ denotes the determinant of the conditional covariance matrix of $(X_0,\ldots,X_{T-1})$ given $(Y_0,\ldots,Y_{T-1})$.[12] Finally, if a variable in the information flow constraint (86) is or may be integrated of order one, then we replace the original variable by its first difference in the information flow constraint to ensure that entropy is always finite.[13]

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 objective (80) is quadratic and the profit-maximizing decisions (82)-(83) follow a Gaussian process.[14] 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 limited attention of individual decision-makers rather than the public 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. We relax this assumption in Section 7.6.

Two remarks are in place before we present solutions of the decision problem (80)-(86). When we solve the decision problem (80)-(86) 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. Furthermore, we evaluate the right-hand side of equation (87) for a very large but finite T.

7.2 Computing the equilibrium of the model

We use an iterative procedure to solve for the rational expectations equilibrium of the model. First, we make a guess concerning the stochastic process for the profit-maximizing price (82) and a guess

[12] If X_t is a vector then Ω_X is the covariance matrix of the vector that is obtained by stacking the vectors $X_0,\ldots,X_{T-1}$ and $\Omega_{X|Y}$ is the conditional covariance matrix of that vector.

[13] If a variable appearing in the information flow constraint (86) follows a stationary Gaussian process, replacing the variable by its first difference in the information flow constraint has no effect on the left-hand side of (86).

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

concerning the stochastic process for the relative wage rate in equation (83). Second, we solve the firms' attention problem (80)-(86). Third, we aggregate the individual prices to obtain the aggregate price level

$$p_t = \frac{1}{I}\sum_{i=1}^{I} p_{it}. \tag{89}$$

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 (57)-(59), equations (61)-(71), 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:

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

$$\tilde{w}_t = \gamma c_t + \psi l_t, \tag{91}$$

$$y_t = c_t, \tag{92}$$

$$y_t = a_t + \alpha l_t, \tag{93}$$

$$a_t = \rho_A a_{t-1} + \varepsilon_t^A, \tag{94}$$

$$r_t = \rho_R r_{t-1} + (1-\rho_R)\left[\phi_\pi\left(p_t - p_{t-1}\right) + \phi_y y_t\right] + \varepsilon_t^R, \tag{95}$$

where 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 (38) and equations (62), (67) and $p_{it} = p_t + \hat{p}_{it}$ imply that the profit-maximizing price is given by

$$p_{it}^* = p_t + \frac{\frac{1-\alpha}{\alpha}}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}} c_t + \frac{1}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\tilde{w}_t - \frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\left(a_t + a_{it}\right). \tag{96}$$

Substituting the law of motion for p_t, c_t, $\tilde{w}_t$, a_t and a_{it} into the last equation yields the law of motion for the profit-maximizing price. In the last equation, we set $\tilde{\theta} = \theta$ because the households' optimality condition (59) and equations (62), (64) and (68) imply that the demand function for good i has the form (19)-(21) with a price elasticity of demand equal to θ. Sixth, if the law of motion for the profit-maximizing price differs from our guess, we update the guess until a fixed point is reached.[15]

Finally, we derive equilibrium relative wage rates. When households have perfect information, equilibrium relative wage rates can be derived analytically. In particular, it is an equilibrium that relative wage rates are constant. The argument is as follows. Suppose that all firms choose the same value for $\tilde{\eta}$ and the same value for χ satisfying $\tilde{\eta} > 1$ and $\chi > 0$. Then, equations (85), (63) and (69) imply that the labor demand function for type j labor has the form (40)-(42) with a wage elasticity of labor demand that is the same for all types of labor. Since all households face the same decision problem and have the same information, all households set the same wage rate. Equation (67) then implies that relative wage rates are constant ($\hat{w}_{jt} = w_{jt} - w_t = 0$). When relative wage rates are constant, the profit-maximizing factor mix is constant, implying that each firm can attain the profit-maximizing factor mix without any information flow. Since each firm can attain the profit-maximizing factor mix without any information flow, no firm has an incentive to deviate from the chosen values for $\tilde{\eta}$ and χ.

7.3 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 of the process for aggregate technology, 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 (93) 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.[16] We then fit equation (94) to the time series for a_t obtaining $\rho_A = 0.96$ and a standard deviation of the innovation equal to 0.0085. In the benchmark economy, we set $\rho_A = 0.95$ and we set the standard deviation of ε_t^A equal to 0.0085.

To set the parameters of the monetary policy rule, we consider quarterly U.S. data on the

[15] We use Matlab and a standard nonlinear optimization program to solve the firms' attention problem. The solution of the firms' attention problem takes about 20 seconds on a machine on which the LU decomposition of a full matrix requires about 0.1 of one second (as reported by the Matlab function *bench.m*). On the way to a fixed point, we make the guess in iteration n a weighted average of the solution in iteration $n-1$ and the guess in iteration $n-1$. The number of iterations needed to reach a fixed point depends significantly on parameter values, on the initial guess, on the weight of the guess in iteration $n-1$ in the guess in iteration n, and on the terminal condition.

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

Federal Funds rate, inflation, and real GDP from 1960 Q1 to 2006 Q4. We fit the monetary policy rule (95) 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.[17] In the benchmark economy, we set $\rho_R = 0.9$, $\phi_\pi = 1.5$, $\phi_y = 0.33$, and we set the standard deviation of ε_t^R equal to 0.0021.

To specify a process for firm-specific productivity, we draw on the recent literature that calibrates menu cost models with firm-specific productivity shocks to micro price data. Nakamura and Steinsson (2008) and Bils, Klenow and Malin (2009) calibrate the autocorrelation of firm-specific productivity to be 0.66 and 0.7, respectively, in monthly models. Since $(0.3)^{1/3}$ equals a number between 0.66 and 0.7, 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 median absolute size of price changes in our model equals 9.7 percent under perfect information. 9.7 percent is the median absolute size of price changes excluding sale-related price changes reported in Klenow and Kryvtsov (2008). This choice yields a standard deviation of the innovation to firm-specific productivity equal to 0.18.[18]

We compute the solution of the model by fixing the marginal cost of information flow, μ. The overall information flow devoted to the price setting and factor mix decision is then determined within the model (i.e. κ is endogenous). See the attention problem (80)-(86). One can interpret the cost μ as an opportunity cost. The idea is that attention devoted to the price setting decision or the factor mix decision could have been devoted to some other decision problem that we do

[17] The specification of the monetary policy rule that we estimate is standard in the empirical literature on the Taylor rule with partial adjustment. See, for example, Section 2 in Rudebusch (2002) for a review of this literature. We regress a measure of the nominal interest rate on its own lag, a measure of the inflation rate, and a measure of the output gap. The nominal interest rate is measured as the quarterly average Federal Funds rate. The inflation rate is measured as $\frac{1}{4}\sum_{l=0}^{3}\pi_{t-l}$, where $\pi_t = \ln P_t - \ln P_{t-1}$ and P_t is the price index for personal consumption expenditures excluding food and energy. The output gap is measured as $(Y_t - Y_t^p)/Y_t^p$, where Y_t is real GDP and Y_t^p is potential real GDP estimated by the Congressional Budget Office. The data sources are the website of the Federal Reserve Bank of St. Louis and the website of the Congressional Budget Office. Note that in the empirical monetary policy rule we measure the inflation rate as the four-quarter moving average of inflation rates. We do so following Section 2 in Rudebusch (2002). Using only the current inflation rate in the empirical monetary policy rule yields essentially identical estimates.

[18] We match the size of price changes excluding sale-related price changes instead of the size of all price changes, because this choice yields a smaller standard deviation of the innovation to firm-specific productivity and therefore less attention will be allocated to firm-specific productivity. For the same reason, we choose the standard deviation of the innovation to firm-specific productivity such that the median absolute size of price changes equals 9.7 percent under perfect information instead of under rational inattention. Again, this choice yields a smaller standard deviation of the innovation to firm-specific productivity.

not model. Alternatively, one can interpret the cost μ as a monetary cost (e.g. a wage payment). We set the marginal cost of information flow equal to 0.1 percent of the firm's revenue in the non-stochastic steady state. We value this marginal cost of information flow in objective (80) using the value of the stochastic discount factor (26) at the non-stochastic steady state. Formally, $\mu = (0.001)\,(1+\tau_p)\,\hat{P}_i Y_i C_j^{-\gamma}$. This value for the marginal cost of information flow will imply that, in equilibrium, the expected per-period loss in profits due to deviations of the price from the profit-maximizing price equals 0.15 percent of the firm's steady state revenue: $(0.0015)\,(1+\tau_p)\,\hat{P}_i Y_i$.[19]

We first report the optimal allocation of attention at the rational inattention fixed point. The decision-maker in a firm who has to set a price devotes most attention to firm-specific productivity, some attention to aggregate technology and little attention to monetary policy. Of the total information flow allocated to the price setting decision, 65 percent is allocated to firm-specific productivity, 26 percent is allocated to aggregate technology, and 9 percent is allocated to monetary policy. Furthermore, for our choice of the marginal cost of information flow, the total attention devoted to the price setting decision is sufficiently high so that firms do very well. The expected per-period loss in profits due to deviations of the price from the profit-maximizing price equals 0.15 percent of the firm's steady state revenue.[20]

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 6 (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. Since the price level does not adjust fully on impact to a monetary policy

[19] To illustrate this number, consider the following simple example. Suppose the firm with a decision-maker with rational inattention has a profit margin of 15 percent. Then, if the decision-maker set the profit-maximizing price in each period, the profit margin of the firm would increase to 15.15 percent. Furthermore, if one part of the decision-maker's compensation is proportional to the profit margin, this part of the decision-maker's compensation would increase by (1/100).

[20] 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 revenue. The expected per-period loss in profits due to imperfect tracking of aggregate technology equals 0.05 percent of the firm's steady state revenue. The expected per-period loss in profits due to imperfect tracking of monetary policy equals 0.03 percent of the firm's steady state revenue.

shock, the real interest rate increases after a positive innovation in the Taylor rule, implying that consumption and output fall. The fall 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 shows a dampened and delayed response to an aggregate technology shock compared with the case of perfect information. For this reason, the output gap is negative for a few quarters after the shock, implying that output shows a hump-shaped impulse response to an aggregate technology shock. Note that the response of the price level to an aggregate technology shock is less dampened and less delayed than the response of the price level to a monetary policy shock. The reason is the optimal allocation of attention. Since decision-makers in firms allocate about three times as much attention to aggregate technology than to monetary policy, prices respond faster to aggregate technology shocks than to monetary policy shocks.[21,22]

Figure 3 shows the impulse response of an individual price to a firm-specific productivity shock. Prices respond very quickly to firm-specific productivity shocks. The reason is that decision-makers in firms decide to pay close attention to firm-specific productivity.

The model can match the empirical finding that the price level responds slowly to monetary policy shocks.[23] Furthermore, the model can match the empirical finding by Altig, Christiano, Eichenbaum and Linde (2005) that the price level responds faster to aggregate technology shocks than to monetary policy shocks. In addition, the model can match the empirical finding by Boivin, Giannoni and Mihov (2009) and Maćkowiak, Moench and Wiederholt (2009) that prices respond very quickly to disaggregate shocks. The reason is the following. We choose the parameter values so as to match key features of the U.S. data like the large average absolute size of price changes

[21] The difference between the response of the price level to monetary policy shocks and the response of the price level to aggregate technology shocks will become even more pronounced once we introduce rational inattention by households.

[22] See also Paciello (2009). Paciello (2009) solves the white noise case of a similar model analytically, where white noise case means: (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.

[23] A number of different identification assumptions lead to the finding that the price level responds slowly to monetary policy shocks. See, for example, Christiano, Eichenbaum and Evans (1999), Leeper, Sims and Zha (1996), and Uhlig (2005).

and the small standard deviation of the innovation in the Taylor rule. For these parameter values, most of the variation in the profit-maximizing price is due to idiosyncratic shocks, a considerable fraction of the variation in the profit-maximizing price is due to aggregate technology shocks, and only a small fraction of the variation in the profit-maximizing price is due to monetary policy shocks. Therefore, the decision-maker in a firm who has to set a price devotes most attention to firm-specific productivity, quite a bit of attention to aggregate technology and little attention to monetary policy. In other words, the decision-maker focuses on those changes in the environment that in expectation cause most of the variation in the profit-maximizing price. In addition, there is an amplification effect. If other firms devote less attention to monetary policy, the profit-maximizing price p_{it}^* given by equation (96) moves less in response to a monetary policy disturbance, which reduces the incentives for an individual firm to devote attention to monetary policy.

7.4 Comparison to the Calvo model

In this subsection, we compare the benchmark economy to the Calvo model with the same preference, technology and monetary policy parameters. We set the Calvo parameter so that prices in the Calvo model change every 2.5 quarters on average. We choose this value for the Calvo parameter because then the impulse responses to a monetary policy shock are essentially identical in the two models. 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). While the impulse responses to a monetary policy shock are essentially identical in the two models, the impulse responses to an aggregate technology shock are very different in the two models. The response of the price level to an aggregate technology shock is much less dampened and delayed in the benchmark economy compared to the Calvo model. As a result, after an aggregate technology shock, the output gap is negative for only 5 quarters in the benchmark economy, whereas the output gap is negative for more than 20 quarters in the Calvo model. Hence, after an aggregate technology shock, the rational inattention model is much closer to a frictionless economy than the Calvo model.

In the benchmark economy and in the Calvo model, firms experience profit losses due to deviations of the price from the profit-maximizing price. It turns out that, in the benchmark economy, the expected loss in profits due to deviations of the price from the profit-maximizing price is about 30 times smaller than in the Calvo model that generates the same response of the price level to monetary policy shocks and the same response of output to monetary policy shocks.[24] The reasons

[24] The expected loss in profits due to suboptimal price responses to aggregate conditions is about 20 times smaller

are the following. 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. By contrast, in the Calvo model, prices respond slowly to all those shocks. Furthermore, under rational inattention deviations of the actual price from the profit-maximizing price are less likely to be extreme than in the Calvo model.

7.5 The effects of changes in parameter values

The model can be used to conduct experiments. Here we report two examples that illustrate how the endogeneity of the allocation of attention affects the outcome of experiments.[25] Furthermore, these two examples show that the outcome of an experiment conducted in the rational inattention model can differ substantially from the outcome of the same experiment conducted in the Calvo model.

First, let us vary the coefficient on inflation in the Taylor rule, ϕ_π. Figure 6 shows the effect of increasing ϕ_π from 1.25 to 1.5 (our benchmark value), to 1.75 and then to 2 on the volatility of the output gap.[26] For comparison, we show the effect in the rational inattention model and the effect in the Calvo model. We report both the standard deviation of the output gap due to monetary policy shocks and the standard deviation of the output gap due to aggregate technology shocks. As ϕ_π increases in the rational inattention model, the standard deviation of the output gap due to monetary policy shocks first falls, bottoming at 1.5, and then rises; and the standard deviation of the output gap due to aggregate technology shocks first rises, peaking at 1.75, and then falls. By contrast, as ϕ_π increases in the Calvo model, the standard deviation of the output gap due to monetary policy shocks declines monotonically and the standard deviation of the output gap due to aggregate technology shocks declines monotonically.

To understand how the value of ϕ_π affects the economy in the two models, note the following. In the Calvo model, as ϕ_π increases the nominal interest rate mimics more closely the real interest rate at the efficient solution. This effect reduces deviations of output from the efficient solution. In the rational inattention model, there is an additional effect. When the central bank responds

than in the Calvo model. The expected loss in profits due to suboptimal price responses to firm-specific conditions is about 40 times smaller than in the Calvo model.

[25] We investigated the role of all parameters in the model. To save space, we only report the effects of changing ϕ_π and α.

[26] Here the output gap is defined as the deviation of aggregate output from equilibrium aggregate output under perfect information. For the economy described in Section 2, it is straightforward to show that due to the subsidies (10)-(11) the equilibrium aggregate output under perfect information equals the efficient aggregate output.

more aggressively with the nominal interest rate to inflation, the price level becomes more stable. Decision-makers in firms react by paying less attention to aggregate conditions. This effect increases deviations of output from the efficient solution. When the second effect dominates the first effect, the volatility of the output gap increases. It turns out that for aggregate technology shocks the second effect dominates for values of ϕ_π below 1.75, while for monetary policy shocks the second effect dominates for values of ϕ_π above 1.5.

Second, consider varying the elasticity of output with respect to composite labor input, α. We find that as α decreases from 1 to 2/3 (our benchmark value) and then toward zero, the price level responds faster to shocks. Again, there are two effects. The first effect is that a decrease in α lowers the coefficient on consumption in the equation for the profit-maximizing price. Formally, substituting equations (91)-(93) and $\tilde{\theta} = \theta$ into equation (96) yields the following equation for the profit-maximizing price

$$p_{it}^* = p_t + \frac{\frac{1-\alpha}{\alpha} + \gamma + \frac{\psi}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\theta} c_t - \frac{\frac{\psi}{\alpha} + \frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\theta} a_t - \frac{\frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\theta} a_{it}. \tag{97}$$

In the benchmark economy, a decrease in α lowers the coefficient on consumption in equation (97).[27] In the language of Ball and Romer (1990), a decrease in α raises the degree of real rigidity, or in the language of Woodford (2003), a decrease in α raises the degree of strategic complementarity in price setting. This effect slows down the adjustment of the price level to shocks. This effect is emphasized in the literature on the Calvo model. The same effect operates here. However, in the rational inattention model, there is an additional effect. As α decreases, the cost of a price setting mistake of a given size increases. Formally, the upper-left element of the matrix H in Proposition 1 increases in absolute value. Thus, decision-makers in firms pay more attention to the price setting decision. This effect speeds up the adjustment of the price level to shocks. We find that the second effect (more attention) dominates the first effect (higher degree of real rigidity) for all values of α between zero and one. Hence, as α decreases, the price level responds faster to shocks.

7.6 Extension: Signals

In this section, we state the attention problem of the decision-maker in a firm using signals. Furthermore, we relax the assumption that attending to aggregate technology, attending to monetary policy and attending to firm-specific productivity are independent activities.

[27] A decrease in α lowers the coefficient on consumption in equation (97) if and only if $\theta(\gamma + \psi) > (1 + \psi)$, which is a parameter restriction that is satisfied in the benchmark economy.

We now assume that, in period -1, the decision-maker in a firm chooses the precision of the signals that he or she will receive in the following periods. In each period $t \geq 0$, the decision-maker receives the signals and takes the optimal price setting and factor mix decision given the signals. The decision-maker chooses the precision of the signals in period -1 so as to maximize the expected discounted sum of profits net of the cost of information flow. The decision-maker understands that a more precise signal (more attention) will lead to better decision making but will also use up more of the valuable information flow. Formally, the attention problem of the decision-maker in firm i reads:

$$\max_{(\kappa,\sigma_1,\sigma_2,\sigma_3,\sigma_4)\in\mathbf{R}_+^5} \left\{ \sum_{t=0}^{\infty} \beta^t E_{i,-1} \left[\frac{1}{2} \left(x_t - x_t^*\right)' H \left(x_t - x_t^*\right) \right] - \frac{\mu}{1-\beta}\kappa \right\}, \tag{98}$$

where

$$x_t - x_t^* = \begin{pmatrix} p_{it} \\ \hat{l}_{i1t} \\ \vdots \\ \hat{l}_{i(J-1)t} \end{pmatrix} - \begin{pmatrix} p_{it}^* \\ \hat{l}_{i1t}^* \\ \vdots \\ \hat{l}_{i(J-1)t}^* \end{pmatrix}, \tag{99}$$

subject to equations (82)-(83) characterizing the profit-maximizing decisions, the following equation characterizing the optimal decision vector in period t given information in period t

$$x_t = E\left[x_t^* | \mathcal{F}_{i0}, s_{i1}, s_{i2}, \ldots, s_{it}\right], \tag{100}$$

the following equation characterizing the signal vector in period t

$$s_{it} = \begin{pmatrix} p_{it}^{A*} \\ p_{it}^{R*} \\ p_{it}^{I*} \\ \hat{w}_{1t} \\ \vdots \\ \hat{w}_{(J-1)t} \end{pmatrix} + \begin{pmatrix} \sigma_1 \nu_{it}^A \\ \sigma_2 \nu_{it}^R \\ \sigma_3 \nu_{it}^I \\ \sigma_4 \nu_{i1t}^L \\ \vdots \\ \sigma_4 \nu_{i(J-1)t}^L \end{pmatrix}, \tag{101}$$

and the constraint on information flow

$$\mathcal{I}\left(\left\{p_{it}^{A*}, p_{it}^{R*}, p_{it}^{I*}, \hat{l}_{i1t}^*, \ldots, \hat{l}_{i(J-1)t}^*\right\}; \{s_{it}\}\right) \leq \kappa. \tag{102}$$

The noise terms ν_{it}^A, ν_{it}^R, ν_{it}^I and ν_{i1t}^L to $\nu_{i(J-1)t}^L$ in the signal 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. As in the decision problem (80)-(86), $E_{i,-1}$ in objective (98) denotes the expectation operator conditioned on the information of the decision-maker in firm i in period -1, the parameter $\mu \geq 0$ in objective (98) is the marginal cost of information flow, and the operator $\mathcal{I}$ in constraint (102) is defined by equation (87). We assume that $E_{i,-1}$ is the unconditional expectation operator. Finally, $\mathcal{F}_{i0}$ in equation (100) denotes the information set of the decision-maker in firm i in period zero. To abstract from transitional dynamics in conditional second moments, we assume that in period zero (i.e. after the decision-maker has chosen the precision of the signals in period -1), the decision-maker receives information such that the conditional variance-covariance matrix of x_t^* given information in period t is constant for all $t \geq 0$.

We solve the problem (98)-(102) for an individual firm, assuming that the aggregate variables are given by the equilibrium of the benchmark economy presented in Section 7.3 and that all relative wage rates are constant. In other words, we assume that the behavior of all other firms and all households is given by the benchmark economy presented in Section 7.3. We then compare the solution to problem (98)-(102) to the solution to problem (80)-(86). Consider Figure 7. The blue lines with points show the impulse responses of the profit-maximizing price to the three fundamental shocks. The green lines with circles show the impulse responses of the price set by the firm to the three fundamental shocks when the firm solves problem (80)-(86). The red lines with crosses show the impulse responses of the price set by the firm to the three fundamental shocks when the firm solves problem (98)-(102). The green lines with circles and the red lines with crosses are identical. Furthermore, the impulse responses of the price set by the firm to the noise terms in equation (84) and to the noise terms in equation (101) also turn out to be identical. In summary, the decision problem (80)-(86) and the decision problem (98)-(102) yield the same price setting behavior.[28]

The signal vector (101) captures the idea that paying attention to aggregate technology, paying

[28] We solve problem (98)-(102) numerically using Matlab and a standard nonlinear optimization program. We first approximate each of the following four objects by an ARMA(p,q) process where p and q are finite: the component of p_t driven by aggregate technology shocks, the component of p_t driven by monetary policy shocks, the component of c_t driven by aggregate technology shocks, and the component of c_t driven by monetary policy shocks. Then, there exists a state-space representation of the dynamics of the signal (101) with a finite-dimensional state vector. We use the Kalman filter to evaluate objective (98) and constraint (102) for any given choice of the precision of the signals. We employ the program *kfilter.m*, written by Lars Ljungqvist and Thomas J. Sargent, to solve for the conditional variance-covariance matrix of the state vector. Solving the problem (98)-(102) takes about as much time as solving the problem (80)-(86). See Footnote 15. Below we also present solutions of problem (98)-(102) with the signal vector (103) instead of the signal vector (101). Solving that problem turned out to be much more time-consuming. Here we had to evaluate objective (98) and constraint (102) on a grid. Standard nonlinear optimization programs proved unhelpful because numerical inaccuracy in the solution for the conditional variance-covariance matrix of the state vector led to spurious variation in the values of the objective and the constraint.

attention to monetary policy, paying attention to firm-specific productivity and paying attention to relative wage rates are independent activities. We now relax this assumption. We replace the signal vector (101) by the following signal vector[29]

$$s_{it}=\begin{pmatrix} p_t \\ a_t+a_{it} \\ c_{i,t-1} \\ w_{t-1}+l_{i,t-1} \\ \hat{w}_{1t} \\ \vdots \\ \hat{w}_{(J-1)t} \end{pmatrix}+\begin{pmatrix} \sigma_1\nu_{i1t} \\ \sigma_2\nu_{i2t} \\ \sigma_3\nu_{i3t} \\ \sigma_4\nu_{i4t} \\ \sigma_5\nu_{i1t}^L \\ \vdots \\ \sigma_5\nu_{i(J-1)t}^L \end{pmatrix}. \tag{103}$$

By choosing σ_1 to σ_5, the decision-maker decides how much attention to devote to the price level, the firm's total factor productivity, the firm's last period sales, the firm's last period wage bill, and the relative wage rates.[30] The variables in the signal vector (103) are driven by multiple shocks and it is therefore no longer the case that, say, paying attention to aggregate technology and paying attention to monetary policy are independent activities. We find that solving the problem (98)-(102) with the signal vector (103) instead of the signal vector (101) changes the firm's price setting behavior hardly at all.[31] See Figure 8. The price set by the firm responds somewhat slower to monetary policy shocks and somewhat faster to aggregate technology shocks. Overall the red lines with crosses in Figure 8 are very similar to the red lines with crosses in Figure 7. We studied a large number of variations of the signal vector (103) and obtained similar results. First, we added other aggregate variables one by one to the signal vector. We found little or no effect on the price setting behavior because the decision-maker of the firm decided to set the precision of the additional signal to a small number or zero. Second, in the signal vector (103) we replaced last period sales and last period wage bill by current period sales and current period wage bill in the signal vector. The

[29] We maintain the assumption that the noise terms in the signal 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.

[30] We include last period sales and last period wage bill in the signal vector because we do not know how the firm can attend to current period sales and current period wage bill before setting the price for its good. Below, when we do assume that the firm can attend to current period sales and current period wage bill, we mean that the firm can attend to the components of current period sales and current period wage bill that are independent of the own price, that is, $\theta p_t + c_t$ and $w_t + (1/\alpha)(\theta p_t + c_t - a_t - a_{it})$, respectively.

[31] When we replace the signal vector (101) by the signal vector (103), we continue to solve the problem of an individual firm, assuming that the aggregate variables are given by the equilibrium of the benchmark economy presented in Section 7.3 and that all relative wage rates are constant.

price set by the firm then responds somewhat faster to monetary policy shocks and to aggregate technology shocks. Still, the price responds more slowly to monetary policy shocks than to aggregate technology shocks. Third, we added firm-specific demand shocks to the model by modifying the consumption aggregator (2). We kept constant the variance of the firm-specific component of the profit-maximizing price. We split this variance equally between firm-specific productivity shocks and firm-specific demand shocks. We assumed the same persistence in firm-specific productivity and in firm-specific demand. We then solved again the decision problem (98)-(102) with the signal vector (103). We found that adding firm-specific demand shocks had almost no effect on the impulse responses of the price set by the firm to monetary policy shocks, to aggregate technology shocks, and to firm-specific productivity shocks. We obtained impulse responses that were almost identical to the red lines with crosses in Figure 8.[32]

8 Rational inattention by firms and households

So far we have assumed that households have perfect information. We now study the implications of adding rational inattention by households. In Sections 8.1-8.2, we solve the model with rational inattention by households assuming households set real wage rates. In Sections 8.3-8.4, we solve the model with rational inattention by households assuming households set nominal wage rates.

We first solve the model assuming households set real wage rates and households have linear disutility of labor because these two assumptions allow us to study in isolation the implications of rational inattention by households for the intertemporal consumption decision. When households have linear disutility of labor ($\psi = 0$), the intratemporal optimality condition stating that the real wage rate should equal the marginal rate of substitution between consumption and leisure reduces to $\tilde{w}_{jt} = \gamma c_{jt}$. Thus, when households have linear disutility of labor and households set real wage rates, households only need to know their own consumption decision to satisfy this intratemporal optimality condition. Knowing the own consumption decision does not require any information flow. Therefore, when $\psi = 0$ and households set real wage rates, households satisfy this intratemporal optimality condition both under perfect information and under rational inattention.

[32] Hellwig and Venkateswaran (2009) also study a model in which firms set prices in period t based on signals concerning sales and wage bills up to and including period $t-1$. There are several differences. First, in their benchmark model the price level and total factor productivity are not included in the signal vector. Second, in their model the noise in the signal is exogenous, whereas in our model the noise in the signal (103) is chosen optimally subject to the constraint on information flow (102). In other words, they report impulse responses for some exogenously given precision of the signals, whereas we report impulse responses for the optimal precision of the signals.

8.1 The households' attention problem when $\psi = 0$ and households set real wage rates

The attention problem of household j in period -1 reads:

$$\max_{\kappa, B_1(L), B_2(L), C_1(L), C_2(L), \tilde{\theta}, \xi} \left\{ \begin{array}{c} \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[\frac{1}{2} \left(x_t - x_t^* \right)' H_0 \left(x_t - x_t^* \right) + \left(x_t - x_t^* \right)' H_1 \left(x_{t+1} - x_{t+1}^* \right) \right] \\ - \frac{\lambda}{1-\beta} \kappa \end{array} \right\}, \tag{104}$$

where

$$x_t - x_t^* = \begin{pmatrix} \tilde{b}_{jt} \\ \tilde{w}_{jt} \\ \hat{c}_{1jt} \\ \vdots \\ \hat{c}_{I-1jt} \end{pmatrix} - \begin{pmatrix} \tilde{b}_{jt}^* \\ \tilde{w}_{jt}^* \\ \hat{c}_{1jt}^* \\ \vdots \\ \hat{c}_{I-1jt}^* \end{pmatrix}, \tag{105}$$

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

$$\tilde{b}_{jt} - \tilde{b}_{jt}^* = - \sum_{l=0}^{t} \left(\frac{1}{\beta} \right)^l \frac{1}{\omega_B} \left[\left(c_{j,t-l} - c_{j,t-l}^* \right) + \tilde{\eta} \omega_W \left(\tilde{w}_{j,t-l} - \tilde{w}_{j,t-l}^* \right) \right], \tag{106}$$

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

$$c_{jt}^* = \underbrace{A_1 (L) \varepsilon_t^A}_{c_{jt}^{A*}} + \underbrace{A_2 (L) \varepsilon_t^R}_{c_{jt}^{R*}} \tag{107}$$

$$\tilde{w}_{jt}^* = \gamma c_{jt}^* \tag{108}$$

$$\hat{c}_{ijt}^* = -\theta \hat{p}_{it}, \tag{109}$$

the equations characterizing the household's actual decisions

$$c_{jt} = \underbrace{B_1 (L) \varepsilon_t^A + C_1 (L) \nu_{jt}^A}_{c_{jt}^A} + \underbrace{B_2 (L) \varepsilon_t^R + C_2 (L) \nu_{jt}^R}_{c_{jt}^R} \tag{110}$$

$$\tilde{w}_{jt} = \gamma c_{jt} \tag{111}$$

$$\hat{c}_{ijt} = -\tilde{\theta} \left(\hat{p}_{it} + \frac{Var \left(\hat{p}_{it} \right)}{\xi} \nu_{ijt}^I \right), \tag{112}$$

and the constraint on information flow

$$\mathcal{I} \left(\left\{ c_{jt}^{A*}, c_{jt}^{R*}, \hat{c}_{1jt}^*, \ldots, \hat{c}_{I-1jt}^* \right\} ; \left\{ c_{jt}^A, c_{jt}^R, \hat{c}_{1jt}, \ldots, \hat{c}_{I-1jt} \right\} \right) \leq \kappa. \tag{113}$$

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 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) 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. The operator $\mathcal{I}$ is defined in equation (87).

Objective (104) states that the household chooses the level and the allocation of information flow so as to maximize the expected discounted sum of period utility net of the cost of information flow. See Proposition 2.[33] The parameter $\lambda \geq 0$ is the per-period marginal cost of information flow. The variable $\kappa \geq 0$ is the information flow devoted to the intertemporal consumption decision, the intratemporal consumption decision and the wage setting decision.[34]

Equations (107)-(109) characterize the decisions that the same household would take if the household had perfect information in every period $t \geq 0$. After the log-quadratic approximation to the expected discounted sum of period utility, the household's optimal decisions under perfect information are given by equations (57)-(60). See Proposition 2. We guess that c_{jt}^* given by equation (57) has the representation (107) at the equilibrium law of motion for r_t and π_t. The guess will be verified. Equations (58) and (59) reduce to equations (108) and (109) after substituting in equation (64) and $\psi = 0$.

We have assumed that the household chooses a consumption vector and a real wage rate. The deviation of the household's real bond holdings in period t from the real bond holdings the same household would have under perfect information is then given by equation (106). Equation (106) follows from equation (60) and $\tilde{b}_{j,-1} = \tilde{b}_{j,-1}^*$. Equation (106) is needed because the deviation $\tilde{b}_{jt} - \tilde{b}_{jt}^*$ is an argument of objective (104).

Equations (110)-(112) characterize the household's actual decisions. Consider first equation (110). By choosing the lag polynomials $B_1(L)$, $C_1(L)$, $B_2(L)$ and $C_2(L)$, the household chooses the stochastic process for composite consumption. For example, if the household chooses $B_1(L) = A_1(L)$, $C_1(L) = 0$, $B_2(L) = A_2(L)$ and $C_2(L) = 0$, the household decides to take the optimal

[33] Proposition 2 states that, after the log-quadratic approximation to expected lifetime utility and for sequences satisfying conditions (51)-(53), maximizing expected lifetime utility is equivalent to maximizing the expression on the right-hand side of equation (54). When we solve the households' attention problem (104)-(113), we consider only stochastic processes for real bond holdings, the real wage rate and the consumption mix that satisfy conditions (51)-(53). It is important to note that conditions (51)-(53) do not require that the processes for real bond holdings, the real wage rate and the consumption mix are stationary. Conditions (51)-(53) do require that second moments increase less than exponentially in t.

[34] We interpret the cost λ as an opportunity cost. In that interpretation, the household has to devote less attention to some other activity in order to devote more attention to the questions of how much to consume, which goods to consume, and which wage to set.

composite consumption decision in every period. The basic trade-off is the following. Choosing a consumption process that tracks more closely optimal composite consumption under perfect information implies smaller utility losses due to suboptimal intertemporal consumption decisions but requires a larger information flow. Equation (111) states that in every period $t \geq 0$ the household sets the real wage rate equal to the marginal rate of substitution between consumption and leisure. The modeling idea behind equation (111) is that the information contained in the household's current and past consumption decisions is also used in the household's current wage setting decision. More precisely, in Appendix E we show analytically that if the household can choose the stochastic process for the real wage rate $\{\tilde{w}_{jt}\}_{t=0}^{\infty}$ as a time-invariant linear one-sided filter of the stochastic process $\left\{c_{jt}^{A}, c_{jt}^{R}\right\}_{t=0}^{\infty}$, then the optimal filter is equation (111) so long as the household has linear disutility of labor ($\psi = 0$). Consider next equation (112). By choosing the coefficients $\tilde{\theta}$ and ξ, the household chooses the price elasticity of demand and the signal-to-noise ratio in the consumption mix decision. The basic trade-off is the following. Choosing a process for the consumption mix that tracks more closely the optimal consumption mix under perfect information reduces the expected loss in utility due to a suboptimal consumption basket but requires a larger information flow.[35]

The constraint (113) states that actual decisions containing more information about the optimal decisions under perfect information require a larger information flow.

Finally, we have to specify the expectation operator $E_{j,-1}$ in objective (104). We assume that all households have perfect information up to and including period -1 and that the particular realization of shocks up to and including period -1 is that shocks are zero. We make this assumption for two reasons. First, this assumption is consistent with the assumption that all households have the same bond holdings in period -1. See Section 2. Second, this assumption implies that all the discounted second moments in objective (104) are finite even when $(x_t - x_t^*)$ has a unit root. We want to allow for the possibility that $(x_t - x_t^*)$ has a unit root.

When we solve the problem (104)-(113) numerically, we turn this infinite-dimensional problem into a finite-dimensional problem by parameterizing each infinite-order lag polynomial $B_1(L)$, $C_1(L)$, $B_2(L)$ and $C_2(L)$ as a lag-polynomial of an ARMA(p,q) process where p and q are finite.

[35] We put more structure on the consumption mix decision than on the intertemporal consumption decision and the wage setting decision. In particular, in equation (112) we express the consumption mix as a function of relative prices rather than expressing the consumption mix as a function of fundamental shocks. We do this because from equation (112) we derive the demand function for good i and a demand function specifies demand on and off the equilibrium path. By expressing the consumption mix as a function of relative prices we specify household j's relative consumption of good i on and off the equilibrium path.

Furthermore, when we solve the problem (104)-(113) numerically, we evaluate the right-hand side of equation (87) for a very large but finite T.

8.2 Benchmark parameter values and solution

We choose the same parameter values as in the benchmark economy in Section 7.3. We have to choose values for five additional parameters: ω_B, ω_W, $\tilde{\eta}$, I and λ. The parameters ω_B, ω_W, $\tilde{\eta}$, I and λ are: the ratio of real bond holdings to consumption in the non-stochastic steady state, the ratio of real wage income to consumption in the non-stochastic steady state, the wage elasticity of labor demand, the number of consumption goods, and the marginal cost of information flow for a household, respectively. All five parameters appear in objective (104).[36] In addition, the parameters ω_B, ω_W and $\tilde{\eta}$ appear in equation (106) because they affect how a percentage deviation in composite consumption and a percentage deviation in the real wage rate translate into a percentage deviation in real bond holdings.

To set the parameters ω_B and ω_W, we consider data from the Survey of Consumer Finances (SCF) 2007. We pursue the following strategy for choosing values for ω_B and ω_W. First, we want to base our calibration of ω_B and ω_W on data for "typical" U.S. households. For this reason, we compute median nominal net worth, median annual nominal wage income and median annual nominal income for the households in the 40-60 income percentile of the SCF 2007. These three statistics equal \$88400, \$41135 and \$47305, respectively. We base our calibration of ω_B and ω_W on all households in the middle income quintile rather than on a single household because we are interested in three variables (net worth, wage income, and income) and the household that is the median household according to one variable may be a very unusual household according to the other two variables. Second, we calculate a proxy for consumption expenditure since consumption appears in the denominator of ω_B and ω_W but the SCF has only limited data on consumption expenditure.[37] The assumption underlying the calculation is that consumption expenditure equals after-tax nominal income minus nominal savings, where nominal savings are just large enough to keep real wealth constant at an annual inflation rate of 2.5 percent. Specifically, we apply the 2007 Federal Tax Rate Schedule Y-1 ("Married Filing Jointly") to annual nominal income given above and we deduct 2.5 percent of nominal net worth given above. This proxy for annual

[36] More precisely, the parameters ω_B, ω_W, $\tilde{\eta}$ and I appear in the matrices H_0 and H_1. See Proposition 2. Thus, these parameters affect the loss in utility in the case of deviations from the optimal decisions under perfect information.

[37] The SCF only has information on the following consumption categories: food, vehicles, housing, and to a very limited extent education and health care.

consumption expenditure equals \$38782. Third, we divide annual nominal wage income given above by four to obtain quarterly nominal wage income, and we divide our proxy for annual consumption expenditure by four to obtain quarterly consumption expenditure. Fourth, we set ω_W equal to the ratio of quarterly nominal wage income to our proxy for quarterly consumption expenditure: $\omega_W = (10283.75/9695.5) = 1.06$, and we set ω_B equal to the ratio of nominal net worth given above to our proxy for quarterly consumption expenditure: $\omega_B = (88400/9695.5) = 9.12$.

We set the wage elasticity of labor demand to $\tilde{\eta} = 4$. In the version of the model with rational inattention by firms and households, decision-makers on the demand side of each market have rational inattention. Therefore, the price elasticity of demand, $\tilde{\theta}$, will typically differ from the preference parameter θ and the wage elasticity of labor demand, $\tilde{\eta}$, will typically differ from the technology parameter η. Throughout the rest of the paper, we set $\tilde{\theta} = 4$ and $\tilde{\eta} = 4$ and we compute the parameter θ that yields a price elasticity of demand of $\tilde{\theta} = 4$ and we compute the parameter η that yields a wage elasticity of labor demand of $\tilde{\eta} = 4$. Thus, we interpret the empirical evidence on price elasticities of demand in the Industrial Organization literature as coming from data generated by our model. For our parameter values, the households' attention problem presented in Section 8.1 has the property that $\theta = 12$ yields $\tilde{\theta} = 4$.[38]

We set the number of consumption goods to $I = 100$. The number of consumption goods does not affect the responses of the household's composite consumption to shocks. The parameter I only affects the household's choice of $\tilde{\theta}$ and ξ. Put differently, the parameter I only affects the θ that yields $\tilde{\theta} = 4$.

We set the marginal cost of information flow equal to the utility equivalent of 0.1 percent of the household's steady state consumption: $\lambda = (0.001)\, C_j * C_j^{-\gamma}$. This value will imply that, in equilibrium, the expected per-period loss in utility due to deviations of composite consumption and of the real wage rate from the optimal decisions under perfect information equals the utility equivalent of 0.04 percent of the household's steady state consumption.[39]

We first solve the attention problem (104)-(113) assuming that aggregate variables and relative prices are given by the equilibrium of the benchmark economy presented in Section 7.3. In other words, we study the optimal allocation of attention by an individual household when decision-makers in firms have limited attention and all other households have perfect information. Figure 9

[38] A price elasticity of demand of four roughly matches estimates of the price elasticity of demand in the Industrial Organization literature.

[39] To fully compensate the household for the expected loss in utility due to deviations of composite consumption and of the real wage rate from the optimal decisions under perfect information, it is sufficient to give the household 1/2500 of the household's steady state consumption in every period.

shows the impulse responses of the household's composite consumption to a monetary policy shock (upper panel) and to an aggregate technology shock (lower panel). The purple lines with squares are the impulse responses of the household's composite consumption under rational inattention. The green lines with circles show what the household would do if the household had perfect information. For the benchmark parameter values, the solution to the attention problem (104)-(113) has several remarkable features. First, the impulse responses of composite consumption under rational inattention are very different from the impulse responses of composite consumption under perfect information, even though the expected per-period loss in utility due to deviations of composite consumption and of the real wage rate from the optimal decisions under perfect information equals the utility equivalent of only 0.04 percent of the household's steady state consumption.[40] Second, the impulse response of composite consumption to a monetary policy shock under rational inattention is hump-shaped, whereas the impulse response under perfect information is monotonic. Third, after a shock to fundamentals composite consumption under rational inattention differs from composite consumption under perfect information, but in the long run the difference between the two impulse responses goes to zero. Similarly, we find that after a shock to fundamentals real bond holdings under rational inattention differ from real bond holdings under perfect information, but in the long run the difference between the two impulse responses goes to zero.[41]

We investigated the role of all parameters of the problem (104)-(113). Here we report two results that we found particularly interesting. First, consider increasing the inverse of the intertemporal elasticity of substitution. As γ increases from 1 (our benchmark value) to 10, the attention devoted to the intertemporal consumption decision increases by 50 percent. Furthermore, the ratio of the actual response to the optimal response of composite consumption on impact of a monetary policy shock increases from 12 percent for $\gamma = 1$ to 26 percent for $\gamma = 10$. In summary, as γ increases from 1 to 10, the household devotes more attention to the intertemporal consumption decision and composite consumption responds faster to a monetary policy shock. However, note that raising γ from 1 to 10 increases the ratio of the actual response to the optimal response of composite consumption on impact of a monetary policy shock only by a factor of about two. This is because there are two effects working in opposite directions. On the one hand, raising γ increases the utility loss in the case of a given deviation of composite consumption from optimal composite consumption.

[40] See Footnote 39.

[41] We also find that the impulse responses of consumption and real bond holdings under rational inattention to the noise terms in equation (110) go to zero in the long run. In the version of the model where all households solve the problem (104)-(113), this finding implies that neither the cross-sectional variance of consumption nor the cross-sectional variance of real bond holdings diverges to infinity.

This effect increases the attention devoted to the intertemporal consumption decision. On the other hand, raising γ lowers the response of optimal composite consumption to the real interest rate. This effect reduces the attention devoted to the intertemporal consumption decision. For γ between 1 and 10, the first effect dominates, but only slightly.

Second, consider varying ω_B and ω_W. We computed five pairs (ω_B, ω_W) by following the same calibration procedure described above for each of the five income quintiles of the SCF 2007. We solved the problem (104)-(113) with the five different pairs (ω_B, ω_W). We then computed a weighted average of the five different impulse responses of composite consumption to a monetary policy shock, where the weight given to each impulse response is the ratio of the proxy for consumption expenditure for this income quintile to the sum of the proxies for consumption expenditure for all five income quintiles. We found that: (i) households in the higher income quintiles have higher values of ω_B and ω_W, (ii) according to the attention problem (104)-(113) households with higher values of ω_B and ω_W pay more attention to the consumption-saving decision and thus their composite consumption responds faster to a monetary policy shock, and (iii) the weighted average of the five impulse responses of composite consumption to a monetary policy shock is very similar to the impulse response of the benchmark household with $(\omega_B, \omega_W) = (9.12, 1.06)$. More precisely, the ratio of the weighted average of the five impulse responses to the impulse response of the benchmark household lies between 0.97 and 1.01 during the first 20 quarters.

For the benchmark parameter values $\gamma = 1$ and $(\omega_B, \omega_W) = (9.12, 1.06)$, we also solved for the new fixed point when decision-makers in firms solve the problem (80)-(86) and all households solve the problem (104)-(113). To save space, we report the main finding in words. The main difference to the impulse responses reported in Figures 1 and 2 is that the impulse response of output to a monetary policy shock becomes hump-shaped. Below we discuss in more detail the new fixed point when firms and households have limited attention but households set nominal wage rates.

8.3 The households' attention problem when $\psi = 0$ and households set nominal wage rates

In this subsection, we state the households' attention problem when households set nominal wage rates instead of real wage rates. The attention problem of a household that sets the nominal wage rate is identical to the problem (104)-(113) apart from the following changes: (i) we add the following equation stating that the log deviation of the real wage rate from the optimal real wage rate under perfect information equals the log deviation of the nominal wage rate from the optimal nominal wage rate under perfect information

$$\tilde{w}_{jt} - \tilde{w}_{jt}^{*} = w_{jt} - w_{jt}^{*}, \tag{114}$$

(ii) we replace equations (107) and (108) by the following equations characterizing the household's optimal composite consumption and *nominal* wage rate decisions under perfect information

$$\begin{aligned} c_{jt}^{*} &= \frac{1}{\gamma}\left(w_{jt}^{*} - p_t\right) &(115)\\ w_{jt}^{*} &= \underbrace{A_1(L)\,\varepsilon_t^A}_{w_{jt}^{A*}} + \underbrace{A_2(L)\,\varepsilon_t^R}_{w_{jt}^{R*}}, &(116)\end{aligned}$$

(iii) we replace equations (110) and (111) by the following equations characterizing the household's actual composite consumption and *nominal* wage rate decisions

$$\begin{aligned} c_{jt} &= \underbrace{D_1(L)\,w_{jt}^A}_{c_{jt}^A} + \underbrace{D_2(L)\,w_{jt}^R}_{c_{jt}^R} &(117)\\ w_{jt} &= \underbrace{B_1(L)\,\varepsilon_t^A + C_1(L)\,\nu_{jt}^A}_{w_{jt}^A} + \underbrace{B_2(L)\,\varepsilon_t^R + C_2(L)\,\nu_{jt}^R}_{w_{jt}^R}, &(118)\end{aligned}$$

and (iv) we add the infinite-order lag polynomials $D_1(L)$ and $D_2(L)$ to the set of objects that the household can choose in period -1.

Equation (114) follows from the fact that $\tilde{w}_{jt} = w_{jt} - p_t$. Equation (114) is needed because the household now sets the nominal wage rate but objective (104) depends on the log deviation of the real wage rate from the optimal real wage rate under perfect information.

Equations (115)-(116) are very similar to equations (107)-(108). Equation (115) is actually identical to equation (108). There are two differences between equations (115)-(116) and equations (107)-(108). First, equations (107)-(108) characterize the optimal composite consumption and real wage rate under perfect information, while equations (115)-(116) characterize the optimal composite consumption and nominal wage rate under perfect information. In addition, there is a small change in notation. Rather than denoting the optimal response of composite consumption to shocks by $A_1(L)$ and $A_2(L)$, we now denote the optimal response of the nominal wage rate to shocks by $A_1(L)$ and $A_2(L)$.[42]

Equations (117)-(118) characterize the household's actual composite consumption decision and the household's actual nominal wage rate decision. Consider first equation (118). By choosing the

[42] Since $c_{jt}^{*} = (1/\gamma)\,\tilde{w}_{jt}^{*}$, this change in notation can also be described as follows. Rather than denoting the optimal response of $(1/\gamma)$ times the real wage rate to shocks by $A_1(L)$ and $A_2(L)$, we now denote the optimal response of the nominal wage rate to shocks by $A_1(L)$ and $A_2(L)$.

lag polynomials $B_1(L)$, $C_1(L)$, $B_2(L)$ and $C_2(L)$, the household chooses the stochastic process for the nominal wage rate. For example, if the household chooses $B_1(L) = A_1(L)$, $C_1(L) = 0$, $B_2(L) = A_2(L)$ and $C_2(L) = 0$, the household decides to set the optimal nominal wage rate in every period. The basic trade-off is again the following. Choosing a process for the nominal wage rate that tracks more closely the optimal nominal wage rate under perfect information reduces utility losses due to suboptimal wage setting but requires a larger information flow. Consider next equation (117). Equation (117) and the fact that the household can choose the lag polynomials $D_1(L)$ and $D_2(L)$ imply that the household can choose the stochastic process for composite consumption $\{c_{jt}\}_{t=0}^{\infty}$ as a time-invariant linear one-sided filter of the stochastic process $\left\{w_{jt}^A, w_{jt}^R\right\}_{t=0}^{\infty}$. The modeling idea behind equation (117) is that the information contained in the household's current and past nominal wage rate decisions is also used in the household's current composite consumption decision.[43]

When we solve the attention problem (104)-(113) with the modifications (114)-(118) numerically, we turn this infinite-dimensional problem into a finite-dimensional problem by parameterizing each infinite-order lag polynomial $B_1(L)$, $C_1(L)$, $D_1(L)$, $B_2(L)$, $C_2(L)$ and $D_2(L)$ as a lag-polynomial of an ARMA(p,q) process where p and q are finite. Furthermore, we evaluate the right-hand side of equation (87) for a very large but finite T.

8.4 Benchmark parameter values and solution

We choose the same parameter values as in Section 8.2. For example, we set the marginal cost of information flow equal to the utility equivalent of 0.1 percent of the household's steady state consumption. This value for the marginal cost of information flow now implies that, in equilibrium, the expected per-period loss in utility due to deviations of composite consumption and of the nominal wage rate from the optimal decisions under perfect information equals the utility equivalent of 0.06 percent of the household's steady state consumption.[44]

[43] We also solved the attention problem of a household that sets the nominal wage rate assuming that the household chooses a consumption process of the form $c_{jt} = B_1(L)\varepsilon_t^A + C_1(L)\nu_{jt}^A + B_2(L)\varepsilon_t^R + C_2(L)\nu_{jt}^R$ and a wage process of the form $w_{jt} = D_1(L)c_{jt}^A + D_2(L)c_{jt}^R$. We found that this alternative setup yields a lower value of objective (104) than the setup (117)-(118). For this reason, we chose the setup (117)-(118). The idea is that if the household could choose between the setup (117)-(118) and the alternative setup, the household would choose the setup (117)-(118). When the household sets the real wage rate, the two setups yield the same solution.

[44] To fully compensate the household for the expected loss in utility due to deviations of composite consumption and of the nominal wage rate from the optimal decisions under perfect information, it is sufficient to give the household 1/1650 of the household's steady state consumption in every period.

We first solve the attention problem of a household that sets the nominal wage rate assuming that aggregate variables and relative goods prices are given by the equilibrium of the benchmark economy presented in Section 7.3. Figure 10 shows the impulse responses of the household's composite consumption to shocks. Figure 11 shows the impulse responses of the household's nominal wage rate to shocks. The purple lines with squares are the impulse responses under rational inattention. The green lines with circles show what the household would do if the household had perfect information. Comparing Figure 10 to Figure 9 shows that composite consumption of a household that sets the nominal wage rate is closer to utility-maximizing composite consumption than composite consumption of a household that sets the real wage rate. The reason is the following. A household that sets the nominal wage rate instead of the real wage rate pays more attention to aggregate conditions because being unaware of changes in aggregate conditions now causes both deviations from the consumption Euler equation and deviations from the intratemporal optimality condition stating that the real wage rate should equal the marginal rate of substitution between consumption and leisure. Specifically, for our parameter values a household that sets the nominal wage rate pays 2.5 times as much attention to aggregate conditions compared to a household that sets the real wage rate. Nevertheless, Figure 10 is similar to Figure 9 in many respects: (i) the impulse responses under rational inattention are very different from the impulse responses under perfect information, even though the expected per-period loss in utility due to deviations of composite consumption and of the nominal wage rate from the optimal decisions under perfect information equals the utility equivalent of only 0.06 percent of the household's steady state consumption, (ii) the impulse response of composite consumption to a monetary policy shock under rational inattention is hump-shaped, and (iii) after a shock to fundamentals composite consumption under rational inattention differs from composite consumption under perfect information, but in the long run the difference between the two impulse responses goes to zero.[45] Furthermore, Figure 11 shows that rational inattention by households also implies that the impulse responses of the nominal wage rate to shocks are dampened and delayed.

Finally, we compute the new fixed point when decision-makers in firms solve the attention problem (80)-(86) and all households solve the attention problem (104)-(113) with the modifica-

[45] We also find that after a shock to fundamentals real bond holdings under rational inattention differ from real bond holdings under perfect information, but in the long run the difference between the two impulse responses goes to zero. Furthermore, we find that the impulse responses of composite consumption, the nominal wage rate and real bond holdings to the noise terms in equation (118) go to zero in the long run. At the new fixed point presented below where all households have limited attention, the last result will imply that neither the cross-sectional variance of consumption nor the cross-sectional variance of real bond holdings diverges to infinity.

tions (114)-(118). We use the following iterative procedure. First, we make a guess concerning the stochastic process for the profit-maximizing price, p_{it}^*, and a guess concerning the stochastic process for the utility-maximizing composite consumption, c_{jt}^*. Second, we solve the firms' attention problem (80)-(86). Third, we aggregate the individual prices to obtain the aggregate price level. Fourth, the guess concerning the stochastic process for the utility-maximizing composite consumption, the law of motion for the aggregate price level and equation (115) yield a guess concerning the stochastic process for the utility-maximizing nominal wage rate, w_{jt}^*. Fifth, we solve the households' attention problem (104)-(113) with the modifications (114)-(118). Sixth, we aggregate across households to obtain aggregate composite consumption, $c_t = \frac{1}{J}\sum_{j=1}^{J} c_{jt}$, and the nominal wage index, $w_t = \frac{1}{J}\sum_{j=1}^{J} w_{jt}$. Seventh, we compute the law of motion for the nominal interest rate from the monetary policy rule (95) and equation (92); we compute the law of motion for the profit-maximizing price from equation (96); and we compute the law of motion for the utility-maximizing composite consumption from equation (57). If the law of motion for the profit-maximizing price or the law of motion for the utility-maximizing composite consumption differs from our guess, we update the guess until a fixed point is reached.[46]

Figures 12 to 14 show the new fixed point. At the new fixed point, the expected per-period loss in profits of a firm due to deviations of the price from the profit-maximizing price equals 0.13 percent of the firm's steady state revenue. The expected per-period loss in utility of a household due to deviations of composite consumption and of the nominal wage rate from the optimal decisions under perfect information equals the utility equivalent of 0.06 percent of the household's steady state consumption. Figure 12 shows impulse responses of the price level, inflation, aggregate composite consumption, and the nominal interest rate to a monetary policy shock. Furthermore, the left panels of Figure 14 show the impulse responses of the real wage index and the nominal wage index to a monetary policy shock. The purple lines with asterisks are the impulse responses at the new fixed point when decision-makers in firms and households have limited attention. For comparison, the green lines with circles show the impulse responses at the old fixed point when only decision-makers in firms have limited attention. See Section 7.3. The main implications of adding rational inattention by households for the impulse responses to a monetary policy shock are the following. The response of aggregate composite consumption to a monetary policy shock becomes hump-shaped; the response of the price level to a monetary policy shock becomes even more dampened and delayed; the response of inflation to a monetary policy shock becomes even more persistent; the

[46] One iteration on the way to a new fixed point takes about 7 minutes on the machine described in Footnote 15. Note that, at each iteration, we now solve the firms' attention problem and the households' attention problem.

response of the real wage index to a monetary policy shock becomes hump-shaped; and the response of the nominal wage index to a monetary policy shock becomes dampened and delayed. Figure 13 shows impulse responses of the price level, inflation, aggregate composite consumption, and the nominal interest rate to an aggregate technology shock. The right panels of Figure 14 show the impulse responses of the real wage index and the nominal wage index to an aggregate technology shock. The main implication of adding rational inattention by households for the impulse responses to an aggregate technology shock is the following. The equilibrium response of aggregate output to an aggregate technology shock becomes more dampened and delayed relative to the efficient response of aggregate output to an aggregate technology shock. To see this, compare the purple line with asterisks in the lower-left panel of Figure 13 to the blue line with points in the lower-left panel of Figure 2.[47]

9 Conclusion

We develop and solve a DSGE model in which households and decision-makers in firms have limited attention and decide how to allocate attention. We find that impulse responses to aggregate shocks display substantial inertia, despite the fact that profit losses and utility losses due to inattention to aggregate conditions are small. This finding suggests that inertia usually modeled with Calvo price setting, habit formation in consumption, and Calvo wage setting may have a different origin. At the same time, our model stands in stark contrast to standard business cycle models when it comes to the mix of slow and fast adjustment of prices to shocks, profit losses due to deviations of the actual price from the profit-maximizing price, and the outcomes of experiments. All those findings suggest that rational inattention may affect the way economists think about business cycles and monetary policy.

Much work remains ahead. One drawback of the model laid out here is the absence of capital. The next step will be to add capital and evaluate the model's quantitative fit to macroeconomic data.

[47]For the economy described in Section 2, it is straightforward to show that due to the subsidies (10)-(11) the equilibrium production under perfect information equals the efficient production.

A Non-stochastic steady state

In this appendix, we characterize the non-stochastic steady state of the economy described in Section 2. We define a non-stochastic steady state as an equilibrium 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 appendix, P_t denotes the following price index

$$P_t = \left(\sum_{i=1}^{I} P_{it}^{1-\theta} \right)^{\frac{1}{1-\theta}}, \tag{119}$$

and W_t denotes the following wage index

$$W_t = \left(\sum_{j=1}^{J} W_{jt}^{1-\eta} \right)^{\frac{1}{1-\eta}}. \tag{120}$$

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

$$\frac{R}{\Pi} = \frac{1}{\beta}, \tag{121}$$

$$\frac{C_{ij}}{C_j} = \hat{P}_i^{-\theta}, \tag{122}$$

and

$$\tilde{W}_j = \varphi \left(\hat{W}_j^{-\eta} L \right)^{\psi} C_j^{\gamma}. \tag{123}$$

The firms' first-order conditions read

$$\hat{P}_i = \tilde{W} \frac{1}{\alpha} \left(\hat{P}_i^{-\theta} C \right)^{\frac{1}{\alpha}-1}, \tag{124}$$

and

$$\hat{L}_{ij} = \hat{W}_j^{-\eta}. \tag{125}$$

The firms' price setting equation (124) 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 (119), the consumption aggregator (2) and the definition of aggregate composite labor input (14) that

$$\hat{P}_i^{1-\theta} = \left(\frac{C_{ij}}{C_j} \right)^{\frac{\theta-1}{\theta}} = \frac{L_i}{L} = \frac{1}{I}. \tag{126}$$

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 vector 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 (120) and the labor aggregator (5) that

$$\frac{C_j}{C} = \hat{W}_j^{1-\eta} = \hat{L}_{ij}^{\frac{\eta-1}{\eta}} = \frac{1}{J}. \tag{127}$$

One can show that equations (121)-(127), $Y_i = L_i^{\alpha}$ and $Y_i = \hat{P}_i^{-\theta} C$ imply that all variables appearing in equations (121)-(127) are uniquely determined apart from the nominal interest rate, R, and inflation, Π. For ease of exposition, we select $\Pi = 1$. Equation (121) then implies $R = (1/\beta)$. Furthermore, the initial price level, P_{-1}, is not determined. We assume that P_{-1} equals some value $\bar{P}_{-1}$. 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).

B 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

$$x_t' = \begin{pmatrix} \hat{p}_{it} & \hat{l}_{i1t} & \cdots & \hat{l}_{i(J-1)t} \end{pmatrix}. \tag{128}$$

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

$$z_t' = \begin{pmatrix} a_t & a_{it} & c_{1t} & \cdots & c_{Jt} & \tilde{w}_{1t} & \cdots & \tilde{w}_{Jt} \end{pmatrix}. \tag{129}$$

Second, we compute a quadratic approximation to the expected discounted sum of profits (30) 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

$$\begin{aligned} & E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t \tilde{f}\left(x_t, z_t\right)\right] \\ = \; & E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t\left(f\left(0,0\right) + h_x' x_t + h_z' z_t + \frac{1}{2}x_t' H_x x_t + x_t' H_{xz} z_t + \frac{1}{2} z_t' H_z z_t\right)\right], \end{aligned} \tag{130}$$

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 (130) using condition (35). Let v_t denote the following vector

$$v_t' = \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}, \tag{131}$$

and let $v_{m,t}$ denote the mth element of v_t. Condition (35) implies that

$$\sum_{t=0}^{\infty} \beta^t E_{i,-1} \left| f(0,0) + h_x' x_t + h_z' z_t + \frac{1}{2} x_t' H_x x_t + x_t' H_{xz} z_t + \frac{1}{2} z_t' H_z z_t \right| < \infty. \tag{132}$$

It follows that one can rewrite equation (130) as

$$\begin{aligned} & E_{i,-1} \left[\sum_{t=0}^{\infty} \beta^t \tilde{f}(x_t, z_t) \right] \\ = & \sum_{t=0}^{\infty} \beta^t E_{i,-1} \left[f(0,0) + h_x' x_t + h_z' z_t + \frac{1}{2} x_t' H_x x_t + x_t' H_{xz} z_t + \frac{1}{2} z_t' H_z z_t \right]. \end{aligned} \tag{133}$$

See Rao (1973), p. 111. Condition (35) also implies that the infinite sum on the right-hand side of equation (133) 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

$$h_x + H_x x_t^* + H_{xz} z_t = 0. \tag{134}$$

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

$$x_t^* = -H_x^{-1} h_x - H_x^{-1} H_{xz} z_t. \tag{135}$$

Hence, x_t^* is uniquely determined and the vector v_t with $x_t = x_t^*$ satisfies condition (35). Fifth, equation (133) implies that

$$\begin{aligned} & E_{i,-1} \left[\sum_{t=0}^{\infty} \beta^t \tilde{f}(x_t, z_t) \right] - E_{i,-1} \left[\sum_{t=0}^{\infty} \beta^t \tilde{f}(x_t^*, z_t) \right] \\ = & \sum_{t=0}^{\infty} \beta^t E_{i,-1} \left[h_x' (x_t - x_t^*) + \frac{1}{2} x_t' H_x x_t - \frac{1}{2} x_t^{*\prime} H_x x_t^* + (x_t - x_t^*)' H_{xz} z_t \right]. \end{aligned} \tag{136}$$

Using equation (134) to substitute for $H_{xz}z_t$ in the last equation and rearranging yields

$$E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t\tilde{f}\left(x_t,z_t\right)\right]-E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t\tilde{f}\left(x_t^*,z_t\right)\right]=\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}\left(x_t-x_t^*\right)'H_x\left(x_t-x_t^*\right)\right]. \quad (137)$$

Sixth, we compute the vector of first derivatives and the matrices of second derivatives appearing in equations (135) and (137). We obtain

$$h_x=0, \quad (138)$$

$$H_x=-C_j^{-\gamma}\tilde{W}L_i\begin{bmatrix} \frac{\tilde{\theta}}{\alpha}\left(1+\frac{1-\alpha}{\alpha}\tilde{\theta}\right) & 0 & \cdots & \cdots & 0 \\ 0 & \frac{2}{\eta J} & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J} \\ \vdots & \frac{1}{\eta J} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1}{\eta J} \\ 0 & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J} & \frac{2}{\eta J} \end{bmatrix}, \quad (139)$$

and

$$H_{xz}=C_j^{-\gamma}\tilde{W}L_i\begin{bmatrix} -\frac{\tilde{\theta}}{\alpha}\frac{1}{\alpha} & -\frac{\tilde{\theta}}{\alpha}\frac{1}{\alpha} & \frac{\tilde{\theta}}{\alpha}\frac{1-\alpha}{\alpha}\frac{1}{J} & \cdots & \frac{\tilde{\theta}}{\alpha}\frac{1-\alpha}{\alpha}\frac{1}{J} & \frac{\tilde{\theta}}{\alpha}\frac{1}{J} & \cdots & \cdots & \frac{\tilde{\theta}}{\alpha}\frac{1}{J} & \frac{\tilde{\theta}}{\alpha}\frac{1}{J} \\ 0 & 0 & 0 & \cdots & 0 & -\frac{1}{J} & 0 & \cdots & 0 & \frac{1}{J} \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & -\frac{1}{J} & \frac{1}{J} \end{bmatrix}, \quad (140)$$

where we used equation (27) in equations (139)-(140). Seventh, substituting equations (138)-(140) into equation (134) yields the following system of J equations:

$$\hat{p}_{it}^*=\frac{\frac{1-\alpha}{\alpha}}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\left(\frac{1}{J}\sum_{j=1}^{J}c_{jt}\right)+\frac{1}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\left(\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}\right)-\frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\left(a_t+a_{it}\right), \quad (141)$$

and

$$\forall j\neq J:\hat{l}_{ijt}^*+\sum_{k=1}^{J-1}\hat{l}_{ikt}^*=-\eta\left(\tilde{w}_{jt}-\tilde{w}_{Jt}\right). \quad (142)$$

Finally, we rewrite equation (142). Summing equation (142) over all $j\neq J$ yields

$$\sum_{j=1}^{J-1}\hat{l}_{ijt}^*=-\eta\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}+\eta\tilde{w}_{Jt}. \quad (143)$$

Substituting the last equation back into equation (142) yields

$$\forall j\neq J:\hat{l}_{ijt}^*=-\eta\left(\tilde{w}_{jt}-\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}\right). \quad (144)$$

Collecting equations (137), (139), (141) and (144), we arrive at Proposition 1.

C 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 (46) that the household can affect in period t

$$x_t' = \begin{pmatrix} \tilde{b}_{jt} & \tilde{w}_{jt} & \hat{c}_{1jt} & \cdots & \hat{c}_{I-1jt} \end{pmatrix}, \tag{145}$$

and, in each period $t \geq 0$, let z_t denote the vector of all variables appearing in the period utility function (46) that the household takes as given

$$z_t' = \begin{pmatrix} r_{t-1} & \pi_t & \tilde{w}_t & l_t & \tilde{d}_t & \tilde{t}_t & \hat{p}_{1t} & \cdots & \hat{p}_{It} \end{pmatrix}. \tag{146}$$

There is one variable appearing in the period utility function (46) 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

$$x_{-1}' = \begin{pmatrix} \tilde{b}_{j,-1} & 0 & \cdots & 0 \end{pmatrix}, \tag{147}$$

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 (46) 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

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

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 (148) 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 (148) using conditions (51)-(53). For all $t \geq 0$, let v_t denote the following vector

$$v_t' = \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}. \tag{149}$$

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 (53) implies that, for all m and n and for $\tau = 0, 1$,

$$\sum_{t=0}^{\infty} \beta^t E_{j,-1} \left| v_{m,t} v_{n,t+\tau} \right| < \infty. \tag{150}$$

Condition (52) implies that condition (150) also holds for $\tau = -1$. It follows that, for all m and n and for $\tau = 0, 1, -1$,

$$E_{j,-1} \left[\sum_{t=0}^{\infty} \beta^t v_{m,t} v_{n,t+\tau} \right] = \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[v_{m,t} v_{n,t+\tau} \right]. \tag{151}$$

See Rao (1973), p. 111. Furthermore, conditions (52)-(53) imply that the infinite sum on the right-hand side of equation (151) converges to an element in $\mathbb{R}$. Thus, conditions (52)-(53) imply that one can rewrite equation (148) as

$$
\begin{aligned}
& E_{j,-1}\left[\tilde{g}\left(x_{-1}, x_0, z_0, x_1, z_1, x_2, z_2, \ldots\right)\right] \\
= \; & g\left(0,0,0,0,0,0,0,\ldots\right) + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[h_x' x_t\right] + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[h_z' z_t\right] \\
& + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2} x_t' H_{x,-1} x_{t-1}\right] + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2} x_t' H_{x,0} x_t\right] \\
& + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2} x_t' H_{x,1} x_{t+1}\right] + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2} x_t' H_{xz,0} z_t\right] + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2} x_t' H_{xz,1} z_{t+1}\right] \\
& + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2} z_t' H_{z,0} z_t\right] + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2} z_t' H_{zx,-1} x_{t-1}\right] + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2} z_t' H_{zx,0} x_t\right] \\
& + \beta^{-1} E_{j,-1}\left[h_{-1}' x_{-1} + \frac{1}{2} x_{-1}' H_{-1} x_{-1} + \frac{1}{2} x_{-1}' H_{x,1} x_0 + \frac{1}{2} x_{-1}' H_{xz,1} z_0\right], \qquad (152)
\end{aligned}
$$

and that each infinite sum on the right-hand side of equation (152) converges to an element in $\mathbb{R}$. In addition, conditions (51)-(52) ensure that the term in the last line on the right-hand side of equation (152) 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 (152) as

$$
\begin{aligned}
& E_{j,-1}\left[\tilde{g}\left(x_{-1}, x_0, z_0, x_1, z_1, x_2, z_2, \ldots\right)\right] \\
= \; & g\left(0,0,0,0,0,0,0,\ldots\right) + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[h_x' x_t\right] + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[h_z' z_t\right] \\
& + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2} x_t' H_{x,0} x_t\right] + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[x_t' H_{x,1} x_{t+1}\right] + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[x_t' H_{xz,0} z_t\right] \\
& + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[x_t' H_{xz,1} z_{t+1}\right] + \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2} z_t' H_{z,0} z_t\right] \\
& + \beta^{-1} E_{j,-1}\left[h_{-1}' x_{-1} + \frac{1}{2} x_{-1}' H_{-1} x_{-1} + x_{-1}' H_{x,1} x_0 + x_{-1}' H_{xz,1} z_0\right]. \qquad (153)
\end{aligned}
$$

Fourth, we define the process $\{x_t^*\}_{t=-1}^{\infty}$. Let E_t denote the expectation operator conditioned on the entire history of the economy up to and including period t. The process $\{x_t^*\}_{t=-1}^{\infty}$ is defined by the following three requirements: (i) x_{-1}^* is given by equation (147), (ii) in each period $t \geq 0$, x_t^* satisfies

$$
E_t\left[h_x + H_{x,-1} x_{t-1}^* + H_{x,0} x_t^* + H_{x,1} x_{t+1}^* + H_{xz,0} z_t + H_{xz,1} z_{t+1}\right] = 0, \qquad (154)
$$

and (iii) the vector v_t with $x_t = x_t^*$ satisfies conditions (51)-(53). Fifth, we derive a result that we will use below. Multiplying equation (154) 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

$$E_t \left[(x_t - x_t^*)' \left(h_x + H_{x,-1} x_{t-1}^* + H_{x,0} x_t^* + H_{x,1} x_{t+1}^* + H_{xz,0} z_t + H_{xz,1} z_{t+1} \right) \right] = 0. \tag{155}$$

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

$$E_{j,-1} \left[(x_t - x_t^*)' \left(h_x + H_{x,-1} x_{t-1}^* + H_{x,0} x_t^* + H_{x,1} x_{t+1}^* + H_{xz,0} z_t + H_{xz,1} z_{t+1} \right) \right] = 0. \tag{156}$$

Rearranging the last equation yields

$$\begin{aligned} & E_{j,-1} \left[(x_t - x_t^*)' \left(h_x + H_{xz,0} z_t + H_{xz,1} z_{t+1} \right) \right] \\ = \; & -E_{j,-1} \left[(x_t - x_t^*)' \left(H_{x,-1} x_{t-1}^* + H_{x,0} x_t^* + H_{x,1} x_{t+1}^* \right) \right]. \end{aligned} \tag{157}$$

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,

$$\left(E_{j,-1} \left[x_{m,t} x_{n,t+\tau}^* \right] \right)^2 \leq E_{j,-1} \left[x_{m,t}^2 \right] E_{j,-1} \left[x_{n,t+\tau}^{*2} \right]. \tag{158}$$

Conditions (51) and (53) and the definition of the process $\{x_t^*\}_{t=-1}^{\infty}$ 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,

$$\left| E_{j,-1} \left[x_{m,t} x_{n,t+\tau}^* \right] \right| < \delta^t A. \tag{159}$$

It follows that the sequence $\left\{ \sum_{t=0}^{T} \beta^t E_{j,-1} \left[x_{m,t} x_{n,t+\tau}^* \right] \right\}_{T=0}^{\infty}$ 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 (153) that

$$\begin{aligned} & E_{j,-1} \left[\tilde{g} \left(x_{-1}, x_0, z_0, x_1, z_1, x_2, z_2, \ldots \right) \right] - E_{j,-1} \left[\tilde{g} \left(x_{-1}^*, x_0^*, z_0, x_1^*, z_1, x_2^*, z_2, \ldots \right) \right] \\ = \; & \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[\frac{1}{2} x_t' H_{x,0} x_t + x_t' H_{x,1} x_{t+1} - \frac{1}{2} x_t^{*\prime} H_{x,0} x_t^* - x_t^{*\prime} H_{x,1} x_{t+1}^* \right] \\ & + \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[(x_t - x_t^*)' \left(h_x + H_{xz,0} z_t + H_{xz,1} z_{t+1} \right) \right] \\ & + \beta^{-1} E_{j,-1} \left[h_{-1}' x_{-1} + \frac{1}{2} x_{-1}' H_{-1} x_{-1} + x_{-1}' H_{x,1} x_0 + x_{-1}' H_{xz,1} z_0 \right] \\ & - \beta^{-1} E_{j,-1} \left[h_{-1}' x_{-1}^* + \frac{1}{2} x_{-1}^{*\prime} H_{-1} x_{-1}^* + x_{-1}^{*\prime} H_{x,1} x_0^* + x_{-1}^{*\prime} H_{xz,1} z_0 \right]. \end{aligned} \tag{160}$$

Substituting $x^*_{-1} = x_{-1}$ and equation (157) into equation (160) 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

$$\begin{aligned} & E_{j,-1}\left[\tilde{g}\left(x_{-1}, x_0, z_0, x_1, z_1, x_2, z_2, \ldots\right)\right] - E_{j,-1}\left[\tilde{g}\left(x^*_{-1}, x^*_0, z_0, x^*_1, z_1, x^*_2, z_2, \ldots\right)\right] \\ = & \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2}\left(x_t - x^*_t\right)' H_{x,0}\left(x_t - x^*_t\right) + \left(x_t - x^*_t\right)' H_{x,1}\left(x_{t+1} - x^*_{t+1}\right)\right]. \end{aligned} \tag{161}$$

Eighth, we compute the vector of first derivatives and the matrices of second derivatives appearing in equations (154) and (161). We obtain

$$h_x' = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \tag{162}$$

$$H_{x,0} = -C_j^{1-\gamma} \begin{bmatrix} \gamma \omega_B^2 \left(1 + \frac{1}{\beta}\right) & \gamma \omega_B \tilde{\eta} \omega_W & 0 & \cdots & 0 \\ \gamma \omega_B \tilde{\eta} \omega_W & \tilde{\eta} \omega_W \left(\gamma \tilde{\eta} \omega_W + 1 + \tilde{\eta} \psi\right) & 0 & \cdots & 0 \\ 0 & 0 & \frac{2}{\theta I} & \cdots & \frac{1}{\theta I} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \frac{1}{\theta I} & \cdots & \frac{2}{\theta I} \end{bmatrix}, \tag{163}$$

$$H_{x,1} = C_j^{1-\gamma} \begin{bmatrix} \gamma \omega_B^2 & \gamma \omega_B \tilde{\eta} \omega_W & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{164}$$

$$H_{x,-1} = \frac{1}{\beta} H'_{x,1}, \tag{165}$$

$$
H_{xz,0} = C_j^{1-\gamma} \left[\begin{array}{cccccccccc}
\frac{\gamma\omega_B^2}{\beta} & -\frac{\gamma\omega_B^2}{\beta} & \frac{\gamma\omega_B\tilde{\eta}^2\omega_W}{\tilde{\eta}-1} & \frac{\gamma\omega_B\tilde{\eta}\omega_W}{\tilde{\eta}-1} & \gamma\omega_B\omega_D & -\gamma\omega_B\omega_T & \frac{\omega_B(1-\gamma)}{I} & \cdots & \frac{\omega_B(1-\gamma)}{I} & \frac{\omega_B(1-\gamma)}{I} \\
\frac{\gamma\omega_B\tilde{\eta}\omega_W}{\beta} & -\frac{\gamma\omega_B\tilde{\eta}\omega_W}{\beta} & \tilde{\eta}^2\omega_W\left(\frac{\gamma\tilde{\eta}\omega_W}{\tilde{\eta}-1}+\psi\right) & \tilde{\eta}\omega_W\left(\frac{\gamma\tilde{\eta}\omega_W}{\tilde{\eta}-1}+\psi\right) & \gamma\tilde{\eta}\omega_W\omega_D & -\gamma\tilde{\eta}\omega_W\omega_T & \frac{\tilde{\eta}\omega_W(1-\gamma)}{I} & \cdots & \frac{\tilde{\eta}\omega_W(1-\gamma)}{I} & \frac{\tilde{\eta}\omega_W(1-\gamma)}{I} \\
0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{I} & \cdots & 0 & \frac{1}{I} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -\frac{1}{I} & \frac{1}{I}
\end{array} \right], \tag{166}
$$

$$
H_{xz,1} = C_j^{1-\gamma} \left[\begin{array}{cccccccccc}
-\frac{\gamma\omega_B^2}{\beta}+\omega_B & \frac{\gamma\omega_B^2}{\beta}-\omega_B & -\frac{\gamma\omega_B\tilde{\eta}^2\omega_W}{\tilde{\eta}-1} & -\frac{\gamma\omega_B\tilde{\eta}\omega_W}{\tilde{\eta}-1} & -\gamma\omega_B\omega_D & \gamma\omega_B\omega_T & -\frac{\omega_B(1-\gamma)}{I} & \cdots & -\frac{\omega_B(1-\gamma)}{I} & -\frac{\omega_B(1-\gamma)}{I} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0
\end{array} \right]. \tag{167}
$$

Ninth, substituting equations (162)-(167) into equation (154) yields the following system of $1+I$ equations:

$$
c_{jt}^* = E_t\left[-\frac{1}{\gamma}\left(r_t - \pi_{t+1} - \frac{1}{I}\sum_{i=1}^{I}\left(\hat{p}_{it+1}-\hat{p}_{it}\right)\right) + c_{jt+1}^*\right], \tag{168}
$$

$$
\tilde{w}_{jt}^* = \frac{\gamma}{1+\tilde{\eta}\psi}c_{jt}^* + \frac{\psi}{1+\tilde{\eta}\psi}\left(\tilde{\eta}\tilde{w}_t + l_t\right) + \frac{1}{1+\tilde{\eta}\psi}\left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \tag{169}
$$

and

$$
\forall i \neq I : \hat{c}_{ijt}^* + \sum_{k=1}^{I-1}\hat{c}_{kjt}^* = -\theta\left(\hat{p}_{it}-\hat{p}_{It}\right), \tag{170}
$$

where the variable c^*_{jt} is defined by

$$c^*_{jt} = \frac{\omega_B}{\beta}\left(r_{t-1} - \pi_t + \tilde{b}^*_{jt-1}\right) - \omega_B \tilde{b}^*_{jt} + \frac{\tilde{\eta}}{\tilde{\eta}-1}\omega_W\left[(1-\tilde{\eta})\,\tilde{w}^*_{jt} + \tilde{\eta}\tilde{w}_t + l_t\right] + \omega_D \tilde{d}_t - \omega_T \tilde{t}_t - \left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right). \tag{171}$$

Finally, we rewrite equation (170). Summing equation (170) 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 (170) yields

$$\forall i \neq I : \hat{c}^*_{ijt} = -\theta\left(\hat{p}_{it} - \frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right). \tag{172}$$

Collecting equations (161), (163), (164), (168), (169), (171) and (172), we arrive at Proposition 2.

D Solution of the model under perfect information

First, the price setting equation (38) and equations (62), (64), (67) and (12) imply that

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

The wage setting equation (58) and equations (62), (64) and (67) imply that

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

The production function (70) and equations (61), (63), (64) and (12) imply that

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

The equation for aggregate output (61) and equations $y_{it} = c_{it}$, (68), (59), (62) and (64) 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 (72)-(74). Furthermore, the consumption Euler equation (57) and equations (62) and (64) imply that

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

Substituting the solution for c_t into the last equation and solving for the real interest rate yields equation (75). Second, the equation for the optimal consumption mix (59) and equation (64) imply equation (76). Note that combining equations (76), (62) and (68) yields a demand function for good i that has the form (19)-(21) with $\tilde{\theta} = \theta$ and $\vartheta = 1$. The price setting equation (38) and equations (62), (67), (72) and (74) and a price elasticity of demand of $\tilde{\theta} = \theta$ imply equation (77). Third, the equation for the optimal factor mix (39) and equation (67) imply equation (78). Note that combining equations (78), (63) and (69) yields a labor demand function that has the form (40)-(42) with $\tilde{\eta} = \eta$ and $\zeta = 1$. 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 (57)-(60) have a unique solution for consumption that is identical for all households. The wage setting equation (58) then implies that all households set the same wage. It follows from equation (67) that $w_t = w_{jt}$, implying $\hat{w}_{jt} = 0$.

E Equation (111)

Proposition 4 *Consider the decision problem (104)-(113). Replace equation (111) by*

$$\tilde{w}_{jt} = D_1 (L) c_{jt}^A + D_2 (L) c_{jt}^R, \tag{173}$$

where $D_1 (L)$ and $D_2 (L)$ are infinite-order lag polynomials, and add the two lag polynomials $D_1 (L)$ and $D_2 (L)$ to the set of objects that the decision-maker can choose in period -1. If $\psi = 0$ then a solution to this decision problem has to satisfy $D_1 (L) = D_2 (L) = \gamma$.

Proof. First, there are multiple consumption and real wage sequences that yield the same bond sequence because the household can finance extra consumption by lowering the real wage and working more. In particular, increasing consumption in period t from c_{jt} to c'_{jt} and lowering the real wage in period t from $\tilde{w}_{jt}$ to $\tilde{w}'_{jt} = \tilde{w}_{jt} - \frac{1}{\tilde{\eta}\omega_W} \left(c'_{jt} - c_{jt} \right)$ leaves the bond sequence unchanged. This follows from equation (106) and $c_{jt} + \tilde{\eta}\omega_W \tilde{w}_{jt} = c'_{jt} + \tilde{\eta}\omega_W \tilde{w}'_{jt}$. Second, for a given stochastic process for bond holdings $\left\{ \tilde{b}_{jt} \right\}_{t=0}^{\infty}$, the only terms in the infinite sum in objective (104) that depend on the random variable $\tilde{w}_{jt}$ are

$$E_{j,-1} \left[\begin{array}{c} \beta^{t-1} C_j^{1-\gamma} \gamma \omega_B \tilde{\eta} \omega_W \left(\tilde{b}_{jt-1} - \tilde{b}_{jt-1}^* \right) \left(\tilde{w}_{jt} - \tilde{w}_{jt}^* \right) \\ -\beta^t C_j^{1-\gamma} \gamma \omega_B \tilde{\eta} \omega_W \left(\tilde{b}_{jt} - \tilde{b}_{jt}^* \right) \left(\tilde{w}_{jt} - \tilde{w}_{jt}^* \right) \\ -\beta^t \frac{1}{2} C_j^{1-\gamma} \tilde{\eta} \omega_W \left(\gamma \tilde{\eta} \omega_W + 1 + \tilde{\eta} \psi \right) \left(\tilde{w}_{jt} - \tilde{w}_{jt}^* \right)^2 \end{array} \right], \tag{174}$$

where we have used equations (55), (56) and (105). Setting the first derivative of the term inside the expectation operator in expression (174) with respect to $\tilde{w}_{jt}$ equal to zero yields

$$\gamma \left[\beta^{-1} \omega_B \left(\tilde{b}_{jt-1} - \tilde{b}_{jt-1}^* \right) - \omega_B \left(\tilde{b}_{jt} - \tilde{b}_{jt}^* \right) \right] = \left(\gamma \tilde{\eta} \omega_W + 1 + \tilde{\eta} \psi \right) \left(\tilde{w}_{jt} - \tilde{w}_{jt}^* \right). \tag{175}$$

Using equation (106), equation (108) and $\psi = 0$, one can rewrite equation (175) as

$$\tilde{w}_{jt} = \gamma c_{jt}. \tag{176}$$

Furthermore, the second derivative of the term inside the expectation operator in expression (174) with respect to $\tilde{w}_{jt}$ is negative. Thus, if $\psi = 0$ then the pair $(c_{jt}, \tilde{w}_{jt})$ that maximizes the term inside the expectation operator in expression (174) for given bond holdings $\tilde{b}_{jt-1}$ and $\tilde{b}_{jt}$ is the one that satisfies equation (176). Third, take a process $\{c_{jt}, \tilde{w}_{jt}\}_{t=0}^{\infty}$ that has a representation of the form (110) and (173) and that does not have the property $D_1(L) = D_2(L) = \gamma$. Define a new process $\left\{c'_{jt}, \tilde{w}'_{jt}\right\}_{t=0}^{\infty}$ by the following two equations

$$c'_{jt} = c_{jt} + \frac{\tilde{\eta}\omega_W}{1 + \gamma\tilde{\eta}\omega_W}\left(\tilde{w}_{jt} - \gamma c_{jt}\right), \tag{177}$$

$$\tilde{w}'_{jt} = \tilde{w}_{jt} - \frac{1}{1 + \gamma\tilde{\eta}\omega_W}\left(\tilde{w}_{jt} - \gamma c_{jt}\right). \tag{178}$$

This new process has the following five properties: (i) the new process $\left\{c'_{jt}, \tilde{w}'_{jt}\right\}_{t=0}^{\infty}$ has a representation of the form (110) and (173), (ii) in each period $t \geq 0$, $c'_{jt} + \tilde{\eta}\omega_W\tilde{w}'_{jt} = c_{jt} + \tilde{\eta}\omega_W\tilde{w}_{jt}$, (iii) in each period $t \geq 0$, $\tilde{w}'_{jt} = \gamma c'_{jt}$, (iv) in each period $t \geq 0$, $c^{A\prime}_{jt}$ is a function only of current and lagged values of c^{A}_{jt}, and (v) in each period $t \geq 0$, $c^{R\prime}_{jt}$ is a function only of current and lagged values of c^{R}_{jt}. Property (i) implies that the new process $\left\{c'_{jt}, \tilde{w}'_{jt}\right\}_{t=0}^{\infty}$ is a feasible choice. Properties (ii)-(iii) imply that if $\psi = 0$ then the new process $\left\{c'_{jt}, \tilde{w}'_{jt}\right\}_{t=0}^{\infty}$ yields a higher value of the infinite sum in objective (104) than the original process $\{c_{jt}, \tilde{w}_{jt}\}_{t=0}^{\infty}$. Properties (iv)-(v) imply that the new process $\left\{c'_{jt}, \tilde{w}'_{jt}\right\}_{t=0}^{\infty}$ does not yield a higher value of the left-hand side of the information flow constraint (113) than the original process $\{c_{jt}, \tilde{w}_{jt}\}_{t=0}^{\infty}$. It follows that if $\psi = 0$ then the original process $\{c_{jt}, \tilde{w}_{jt}\}_{t=0}^{\infty}$ cannot be a solution to the decision problem specified in Proposition 4. ■

References

[1] Altig, David, Lawrence Christiano, Martin Eichenbaum and Jesper Linde (2005): "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] Ball, Lawrence, and David Romer (1990): "Real Rigidities and the Non-Neutrality of Money." Review of Economic Studies, 57(2), 183-203.

[4] Bils, Mark, Peter J. Klenow and Benjamin A. Malin (2009): "Reset Price Inflation and the Impact of Monetary Policy Shocks." Discussion paper, University of Rochester, Stanford University, and Federal Reserve Board.

[5] Boivin, Jean, Marc Giannoni and Ilian Mihov (2009): "Sticky Prices and Monetary Policy: Evidence from Disaggregated U.S. Data." American Economic Review, 99(1), 350-384.

[6] Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans (1999): "Monetary Policy Shocks: What Have We Learned and to What End?" In "Handbook of Macroeconomics", edited by John B. Taylor and Michael Woodford, Elsevier, New York.

[7] Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." Journal of Political Economy, 113(1), 1-45.

[8] Hellwig, Christian, and Venky Venkateswaran (2009): "Setting the Right Prices for the Wrong Reasons." Journal of Monetary Economics, 56(S), 57-77.

[9] 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.

[10] Leeper, Eric M., Christopher A. Sims, and Tao Zha (1996): "What Does Monetary Policy Do?" Brookings Papers on Economic Activity, 1996(2), 1-63.

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

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

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

[14] Maćkowiak, Bartosz, Emanuel Moench and Mirko Wiederholt (2009): "Sectoral Price Data and Models of Price Setting." Journal of Monetary Economics, 56(S), 78-99.

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

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

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

[18] Paciello, Luigi (2009): "Monetary Policy Activism and Price Responsiveness to Aggregate Shocks under Rational Inattention." Discussion paper, Einaudi Institute for Economics and Finance.

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

[20] Rudebusch, Glenn D. (2002): "Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia." Journal of Monetary Economics 49(6), 1161-1187.

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

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

[23] Smets, Frank, and Rafael Wouters (2007): "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach." American Economic Review, 97(3), 586-606.

[24] Uhlig, Harald (2005): "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure." Journal of Monetary Economics, 52, 381-419.

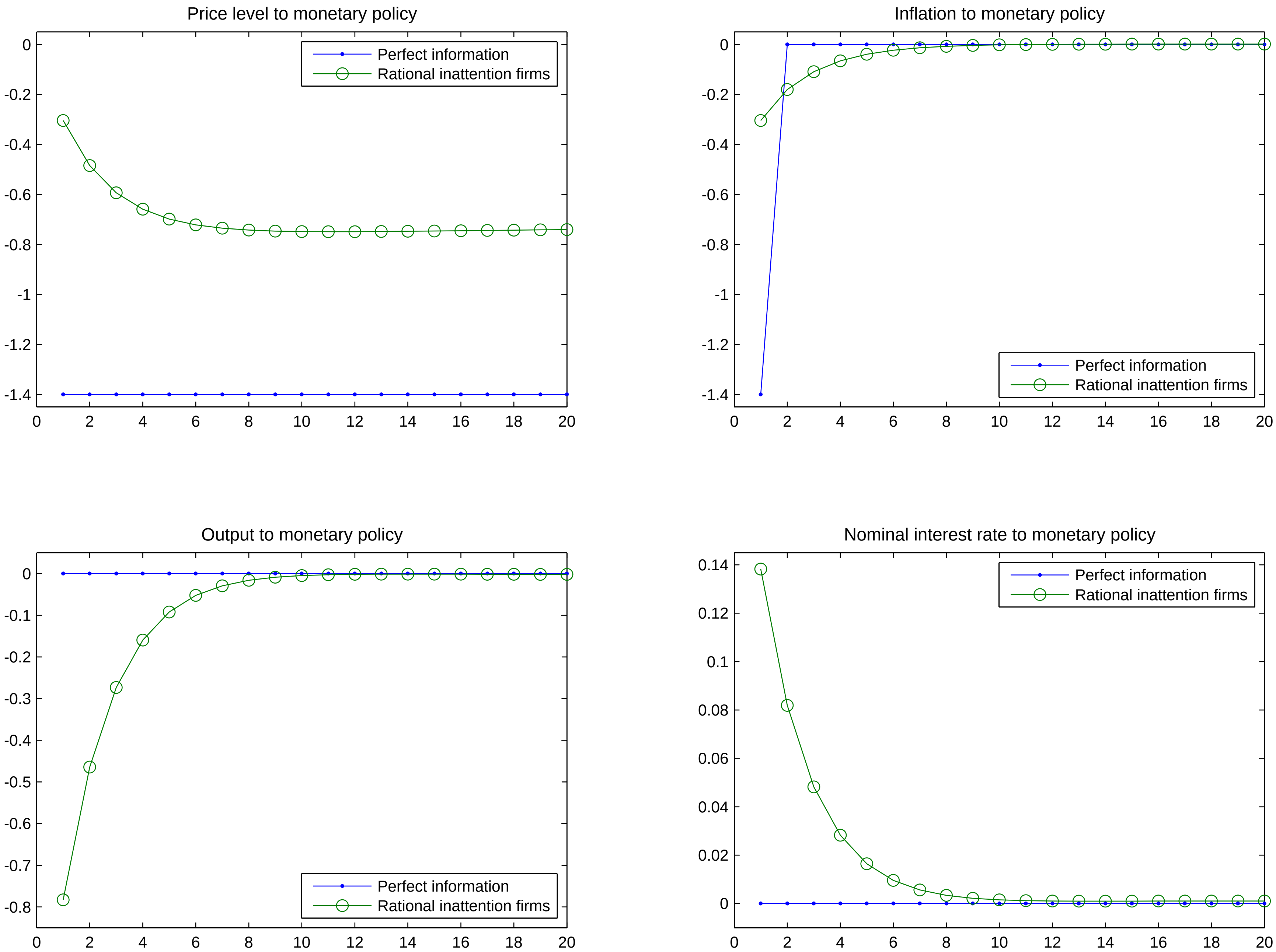
[25] Van Nieuwerburgh, Stijn, and Laura Veldkamp (2009): "Information Acquisition and Under-Diversification." Review of Economic Studies, forthcoming.

[26] 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.

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

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

Figure 1: Impulse responses, benchmark economy



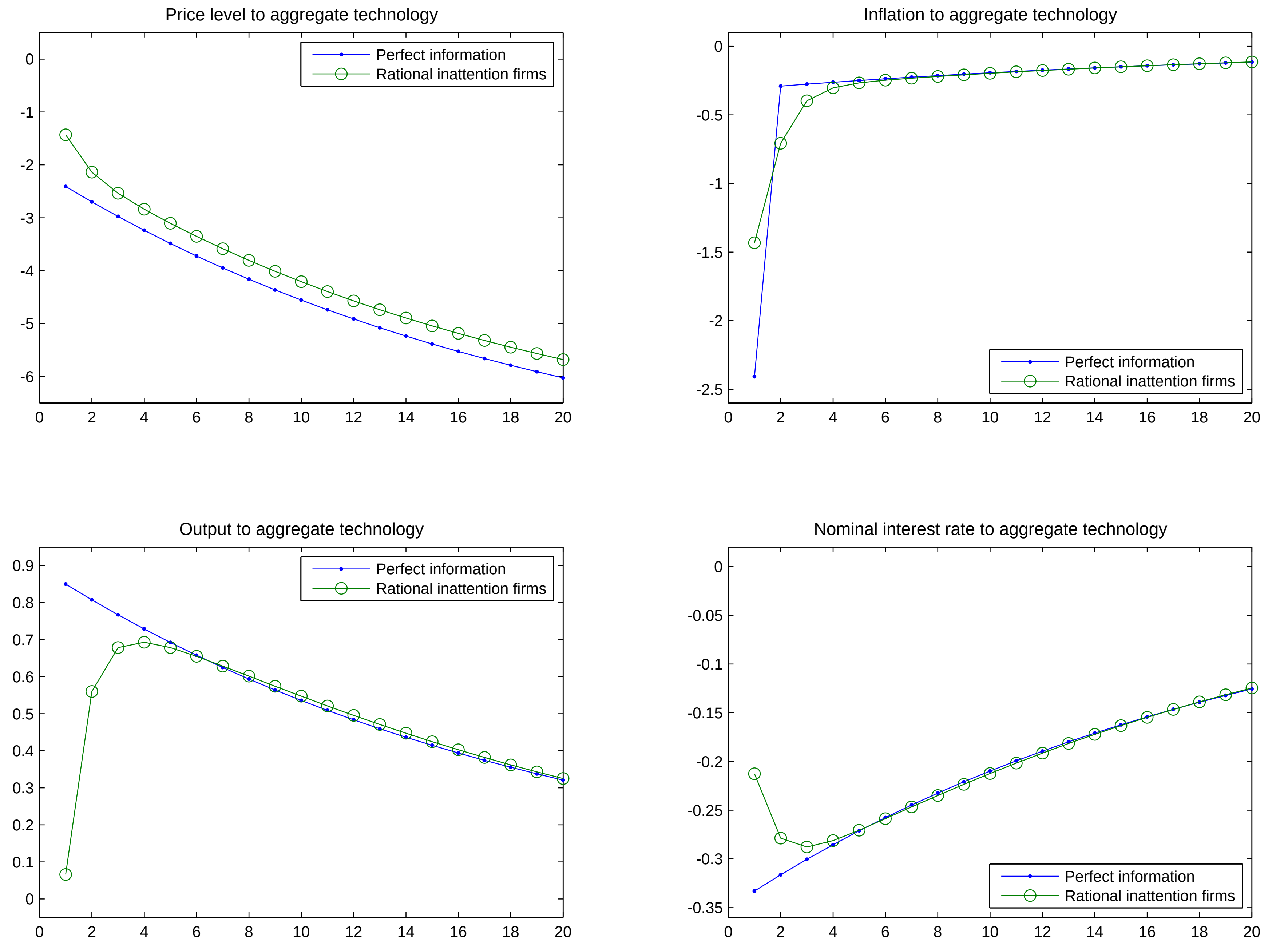


Figure 2: Impulse responses, benchmark economy

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

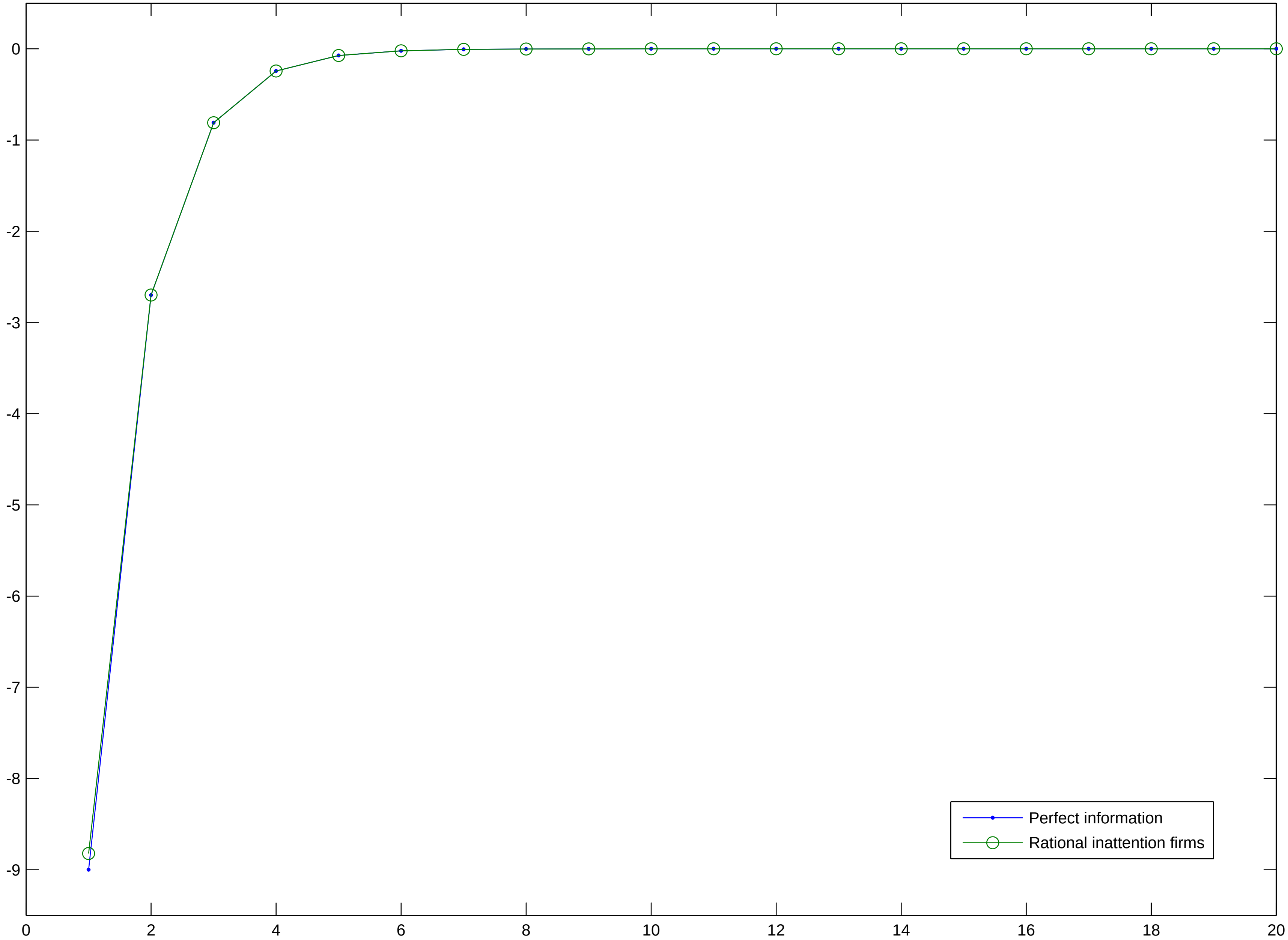


Figure 4: Impulse responses, benchmark economy and Calvo model

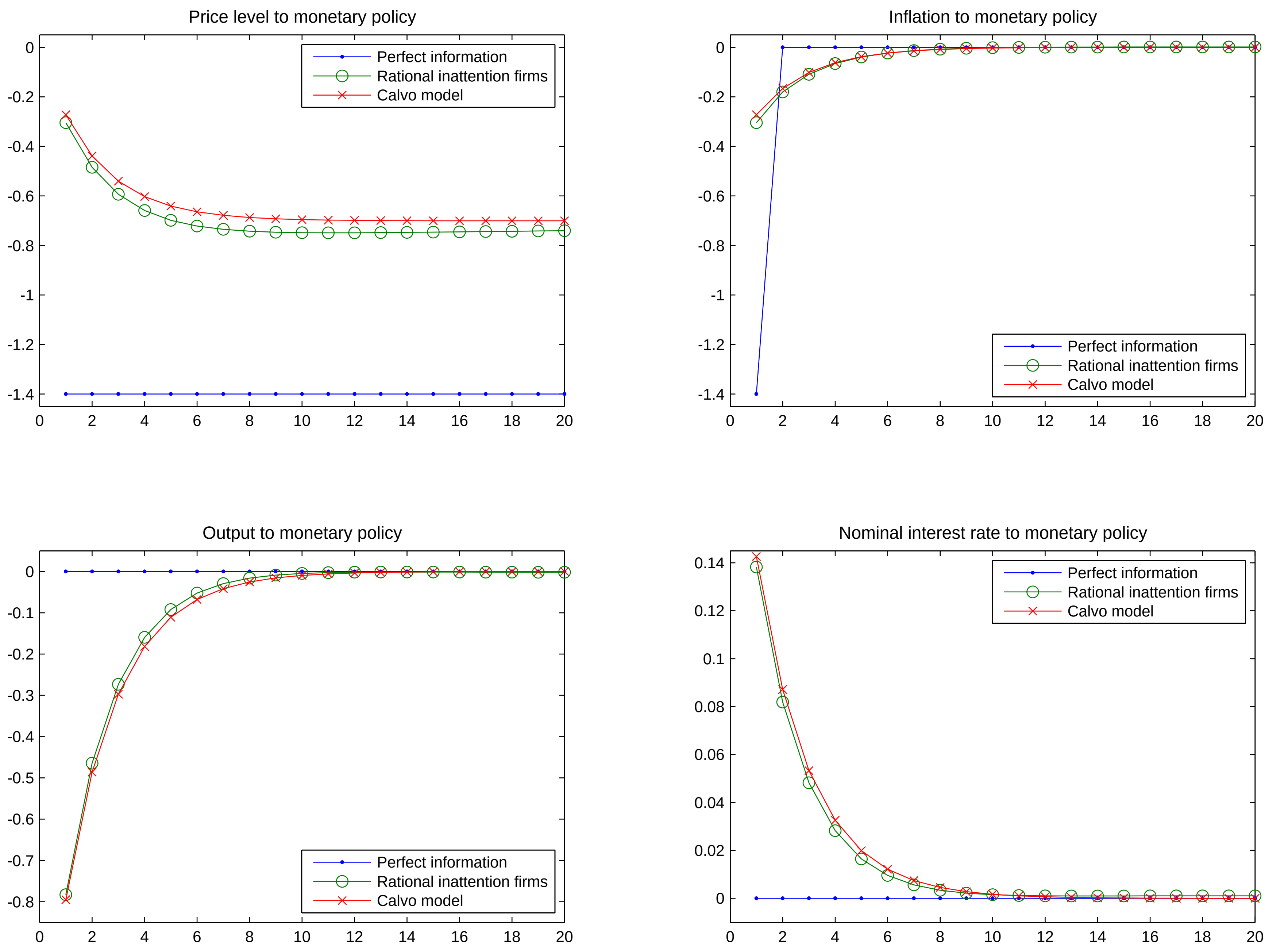
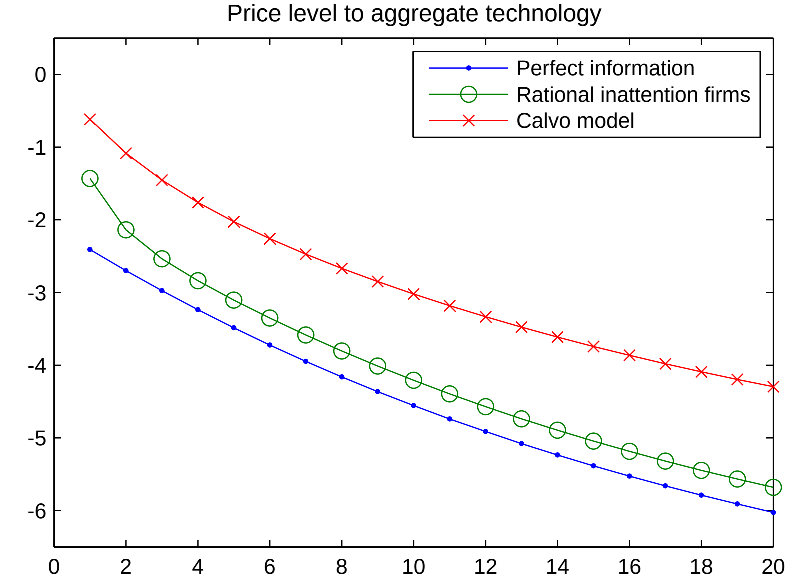
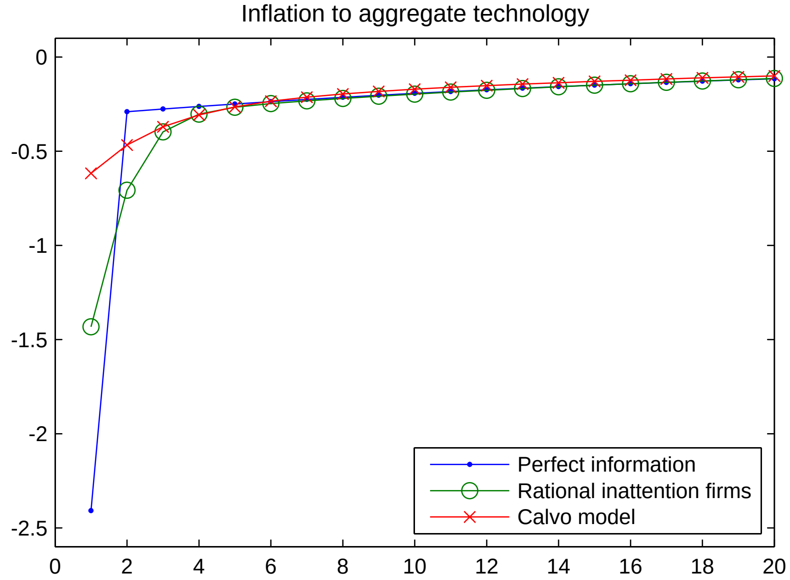
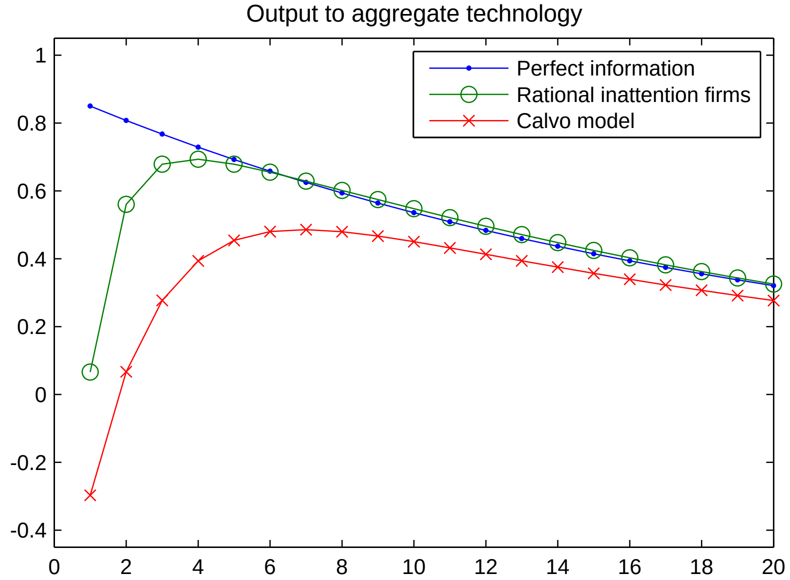


Figure 5: Impulse responses, benchmark economy and Calvo model







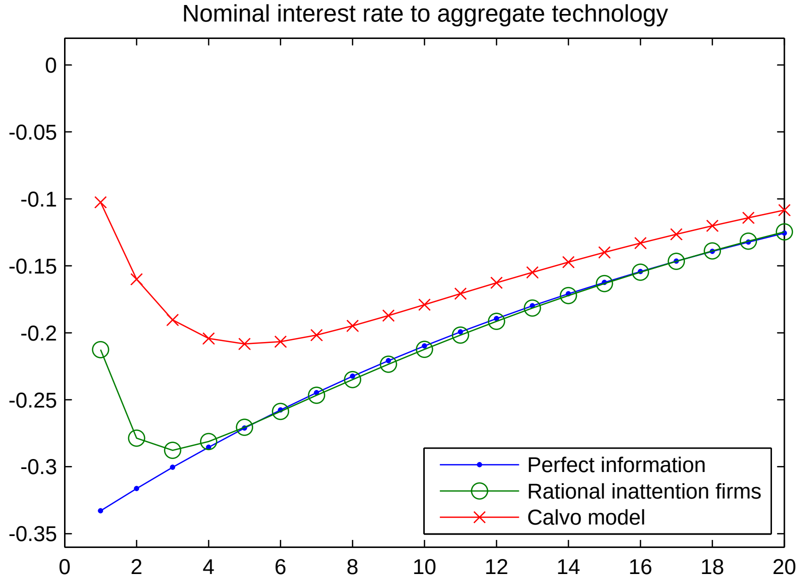


Figure 6: Standard deviation of output gap vs. parameter ϕ_π, benchmark economy and Calvo model

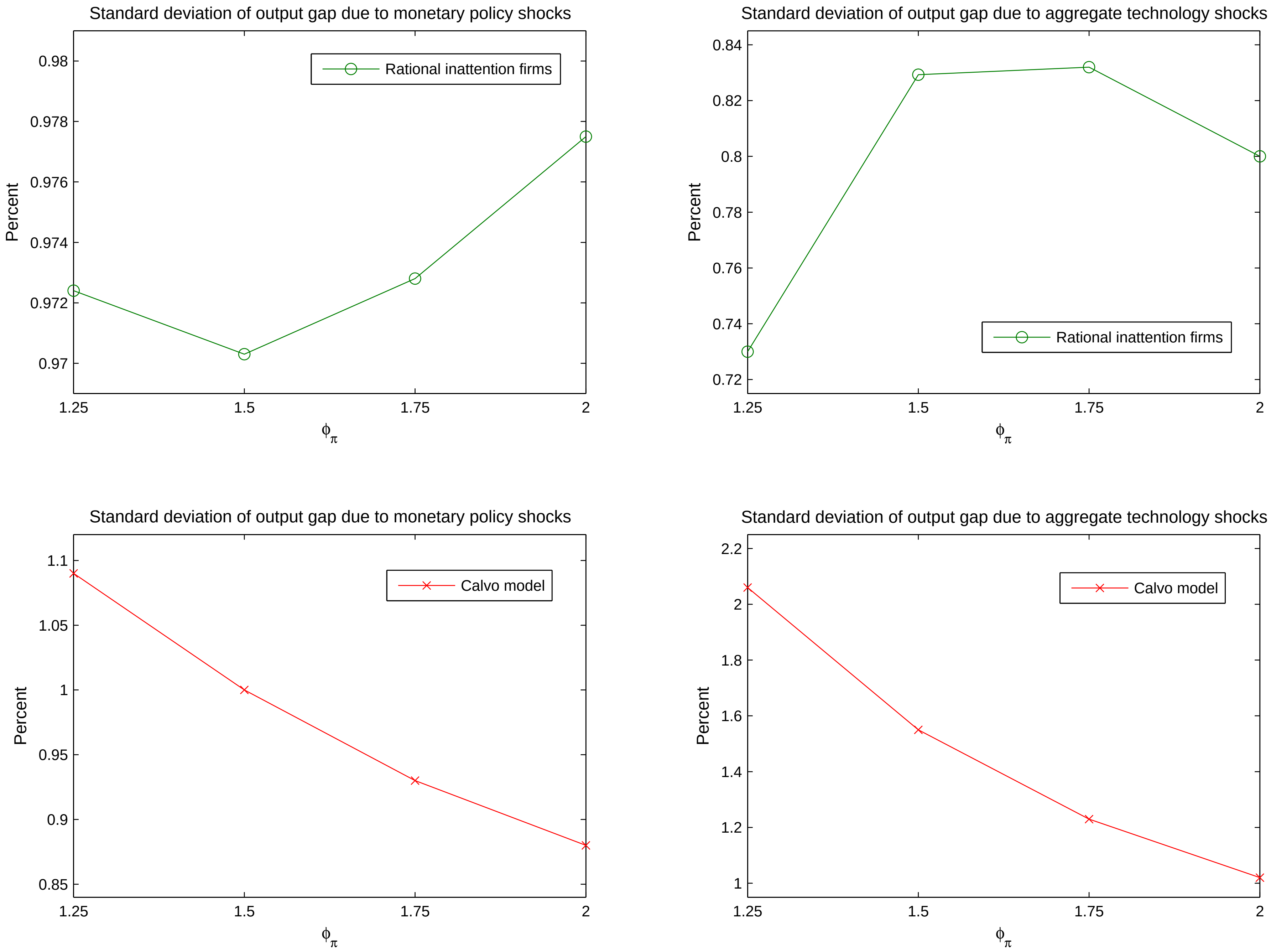


Figure 7: Firms' attention problem with signals concerning p^{A*}, p^{R*}, and p^{I*}

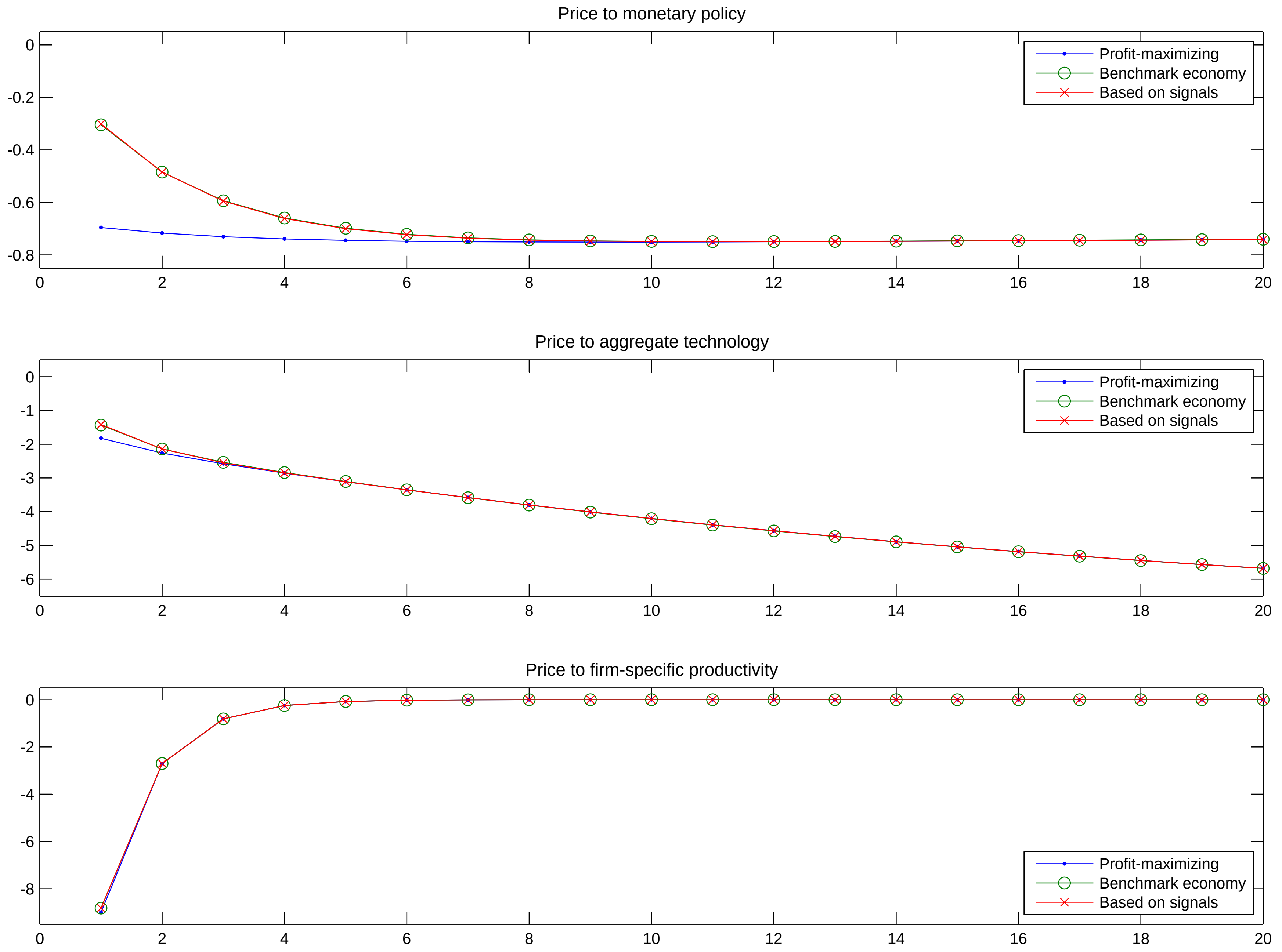


Figure 8: Firms' attention problem with signals concerning the price level, TFP, last period sales, and last period wage bill

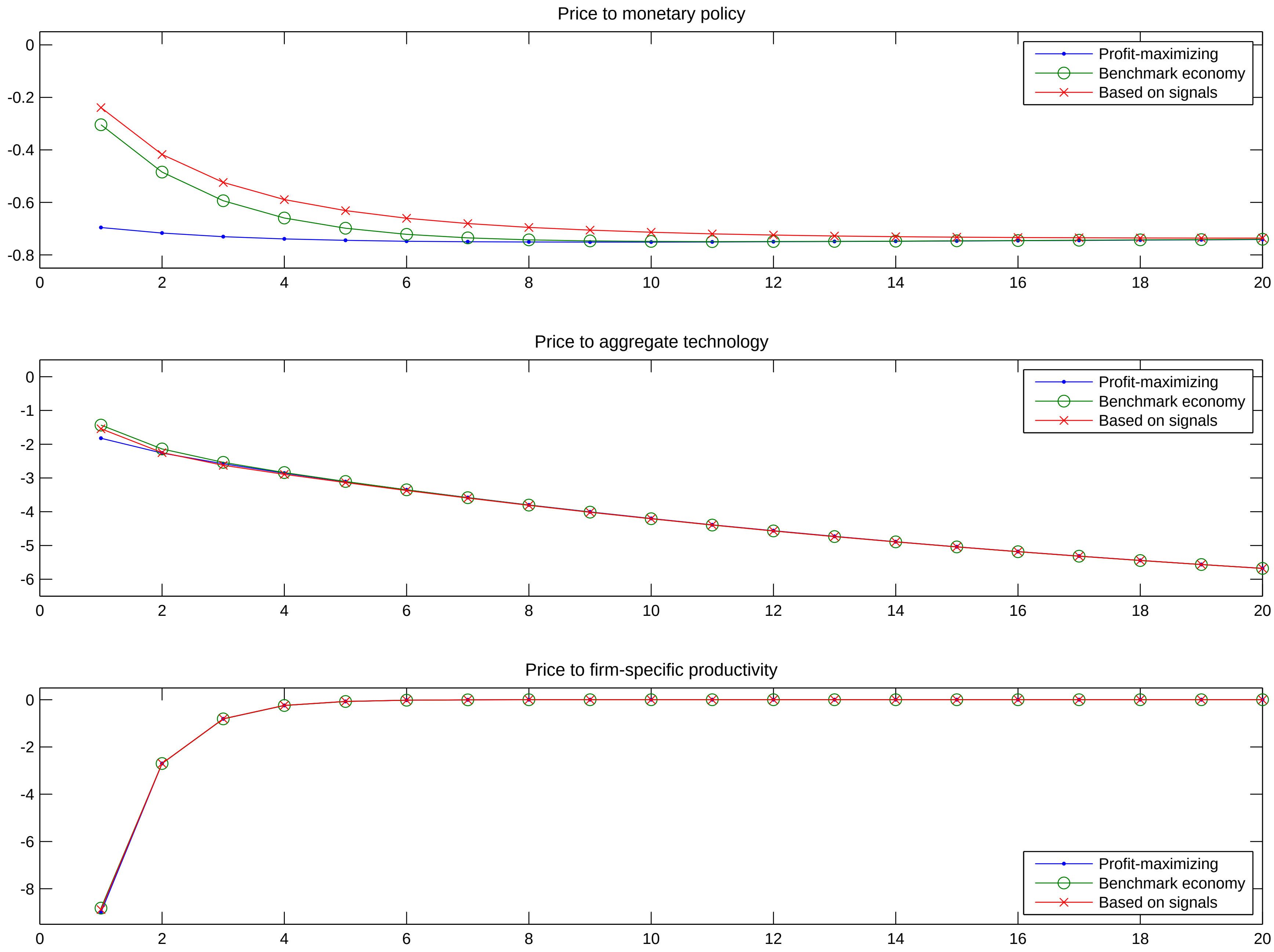


Figure 9: Impulse responses, households' problem

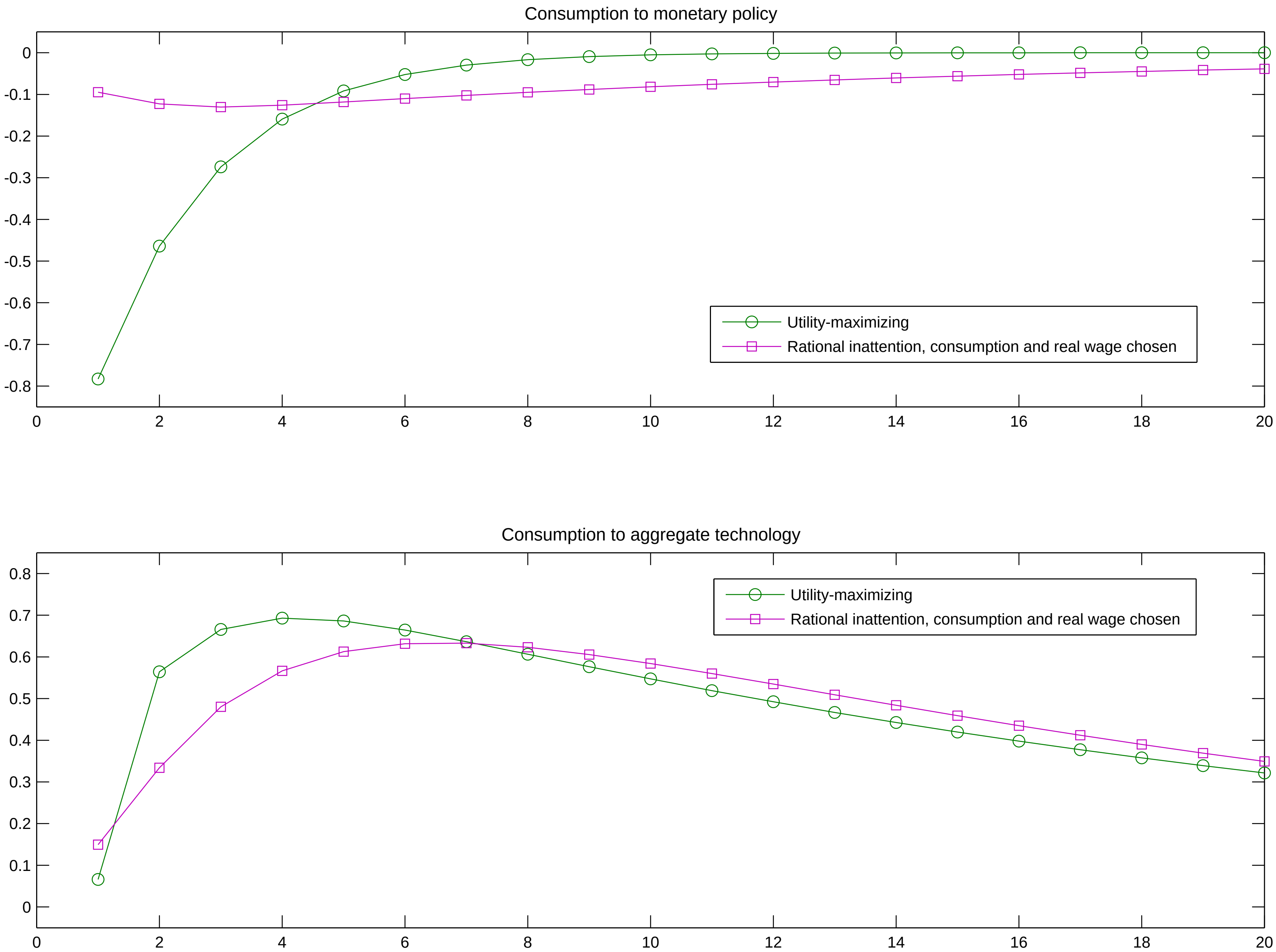
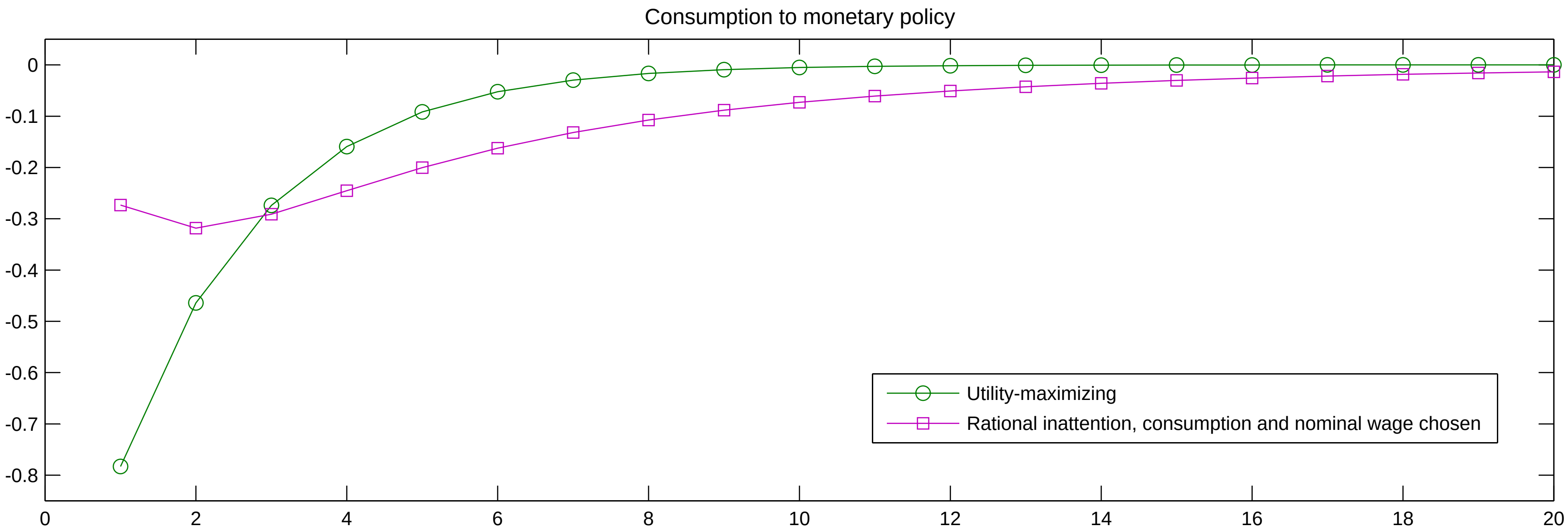


Figure 10: Impulse responses, households' problem



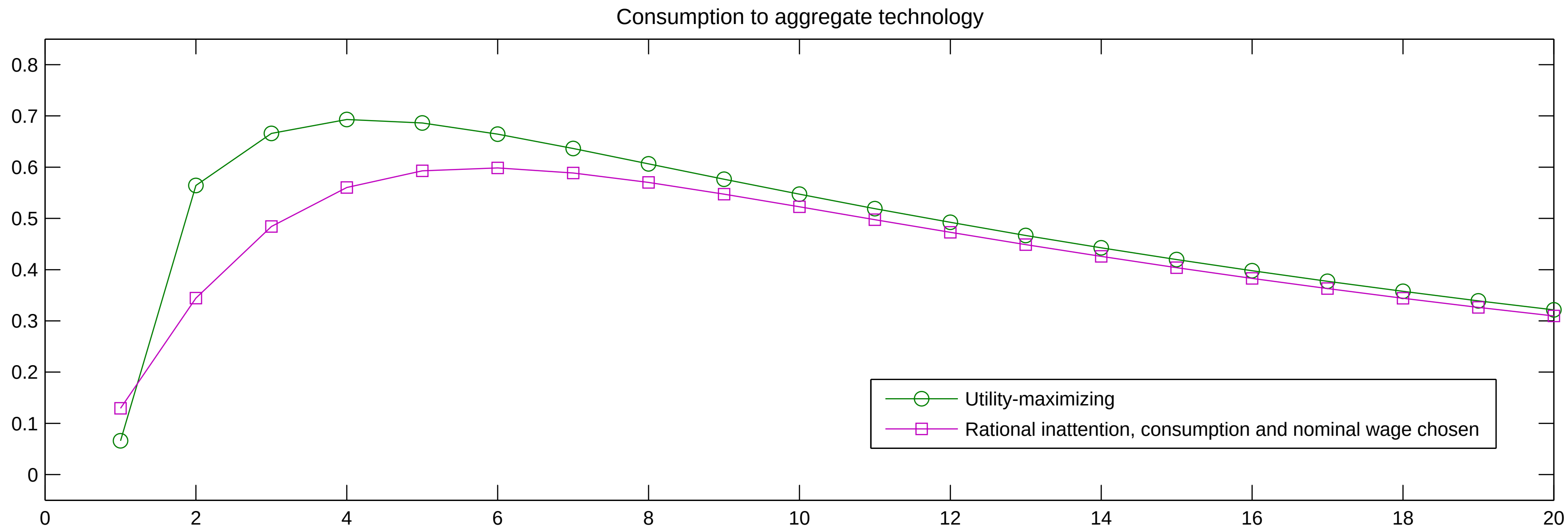


Figure 11: Impulse responses, households' problem

Figure 12: Impulse responses, benchmark economy

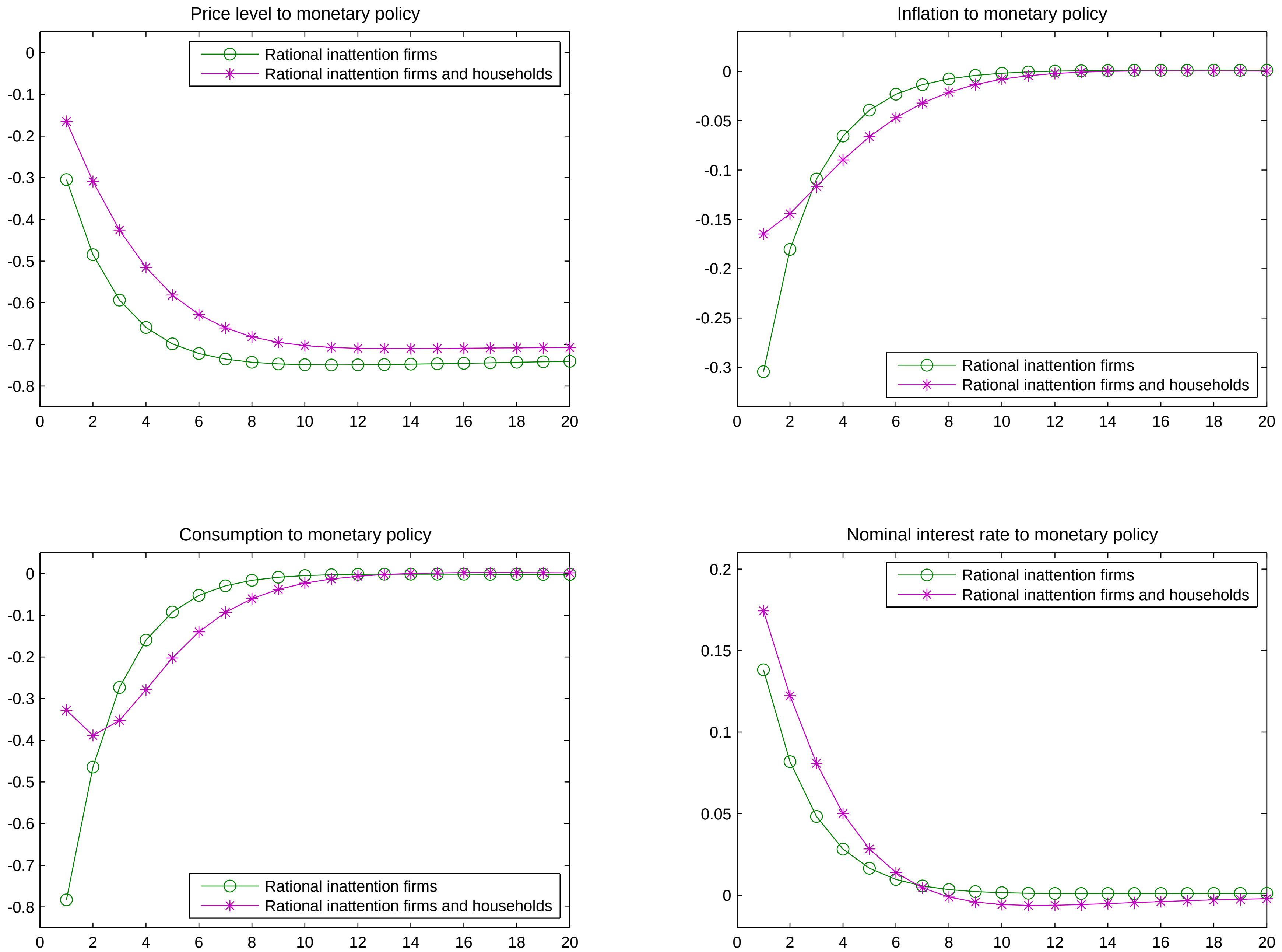


Figure 13: Impulse responses, benchmark economy

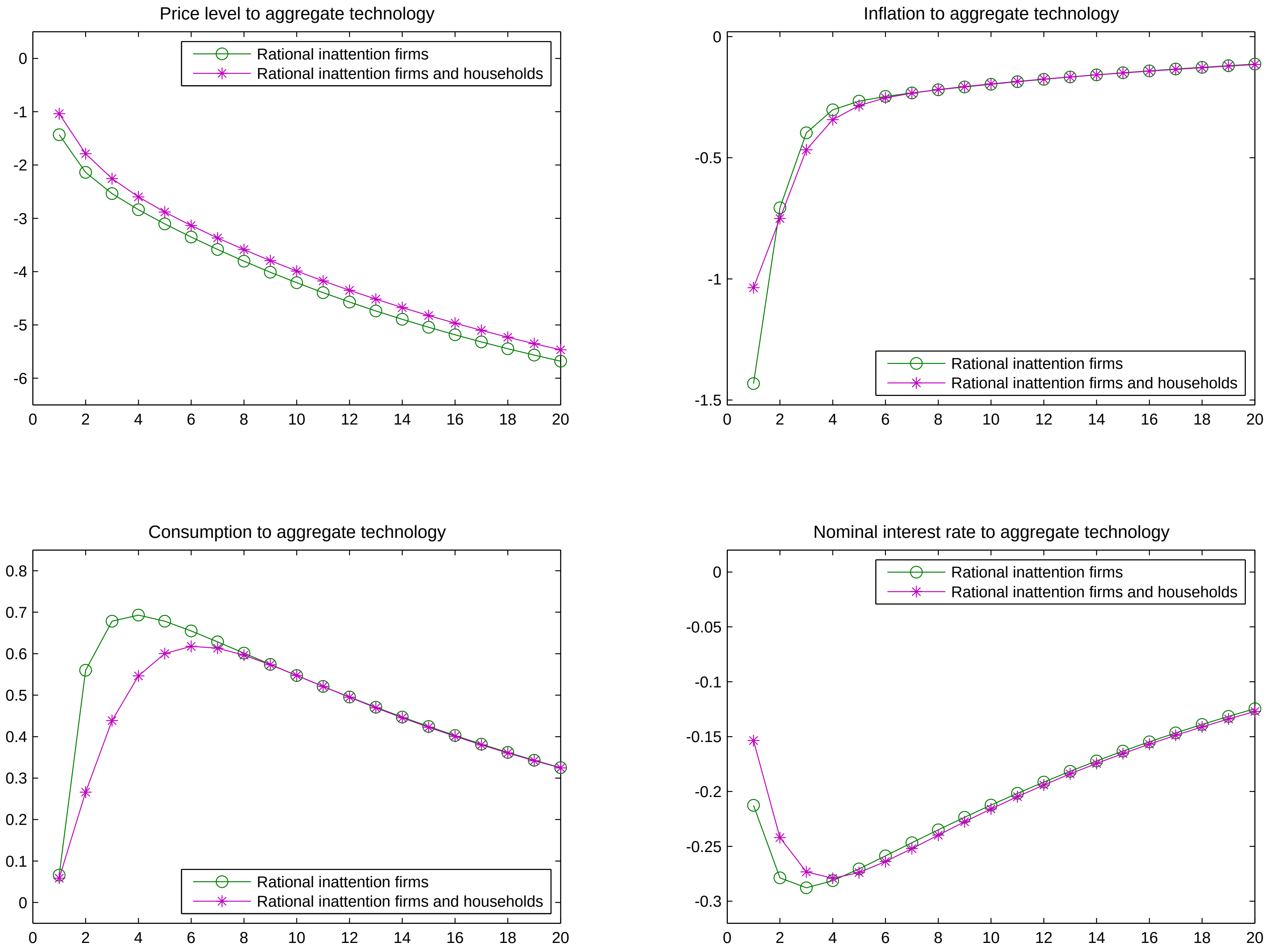


Figure 14: Impulse responses, benchmark economy