Business Cycle Dynamics under Rational Inattention*

Bartosz Maćkowiak
European Central Bank and CEPR

Mirko Wiederholt
Northwestern University

First draft: June 2007. This draft: May 2010

Abstract

This paper develops a dynamic stochastic general equilibrium (DSGE) 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. The model can match empirical findings that are difficult to match with other DSGE models. Moreover, due to the endogeneity of the allocation of attention, the outcomes of policy experiments conducted with this model (e.g., the effects of changing parameters of the monetary policy rule) differ markedly from the outcomes of the same policy experiments conducted with other DSGE models.

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, LSE, Madison, Mannheim, 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

Economists have studied for a long time how decision-makers allocate scarce resources. The recent literature on rational inattention studies how decision-makers allocate the scarce resource attention. The idea is that decision-makers have limited attention and decide how to allocate their attention. This paper develops a dynamic stochastic general equilibrium (DSGE) model with rational inattention. Decision-makers in firms and households have limited attention and decide how to allocate their attention. Following Sims (2003), we model attention as an information flow and we model limited attention as a constraint on information flow. 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 knowing the real interest rate and the prices of all consumption goods requires attention, households have limited attention, and 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 slow adjustment to shocks is rational inattention by decision-makers.

We summarize the model's predictions in four points. The first prediction of the model is that prices respond rapidly to market-specific shocks, fairly quickly to aggregate technology shocks, and slowly to monetary policy shocks. We first solve the model assuming rational inattention by

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

decision-makers in firms and perfect information on the side of households to isolate the implications of rational inattention by decision-makers in firms. We find that: (i) prices respond very quickly to market-specific shocks, (ii) the price level responds fairly quickly to aggregate technology shocks, and (iii) the price level responds slowly to monetary policy shocks. The reason for this combination of fast and slow adjustment of prices to shocks is that decision-makers in firms decide to pay a lot of attention to market-specific conditions, quite a bit of attention to aggregate technology, and little attention to monetary policy. The empirical literature finds in the data the same pattern of fast and slow responses of prices to shocks.[2] This pattern of fast and slow adjustment of prices to shocks is difficult to match with DSGE models that are commonly used for monetary policy analysis (e.g., the Calvo model or the sticky information model of Mankiw and Reis (2002)).

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. An important 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 but fairly quickly to aggregate technology shocks and rapidly to market-specific shocks. By contrast, in the Calvo model prices respond slowly to all 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.

The second prediction of the model is that households with rational inattention respond very slowly with their consumption-saving decision to movements in the real interest rate. When we solve the model with rational inattention by decision-makers in firms and rational inattention by households, we find that households decide to pay little attention to movements in the real interest rate. This finding is important because in a large class of models monetary policy affects the real economy through the effect of the real interest rate on consumption. Moreover, the finding that households decide to pay little attention to movements in the real interest rate turns out to hold for low and high values of the coefficient of relative risk aversion. The reason is the following. For low

[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.

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 and therefore households decide to pay little attention to movements in the real interest rate. As a result, consumption responds very slowly to monetary policy shocks.

The third set of predictions of the model concern how firms and households interact in general equilibrium under rational inattention. When we solve the model under rational inattention by decision-makers in firms and households, we find that adding rational inattention by households has the following implications for aggregate dynamics. First, the impulse response of aggregate consumption to a monetary policy shock becomes hump-shaped. The reason is that households decide to pay little attention to movements in the real interest rate and therefore respond slowly with their consumption to monetary policy shocks. Second, the impulse response of the price level to a monetary policy shock becomes even more dampened and delayed, compared to the case with rational inattention by decision-makers in firms only. The main reason is that households' optimal allocation of attention affects firms' optimal allocation of attention. The dampened and delayed response of aggregate consumption to monetary policy shocks makes decision-makers in firms pay even less attention to monetary policy, implying that the price level responds even more slowly to monetary policy shocks. Third, for the same reasons, adding rational inattention by households also implies that the responses of aggregate consumption and the price level to an aggregate technology shock become more dampened and delayed.

One can compare the DSGE model developed here to the DSGE models commonly used for monetary policy analysis (i.e., the standard New Keynesian model and the sticky information model). The model developed here can match empirical findings that those models can match as well (e.g., the slow response of the price level to monetary policy shocks and the hump-shaped response of consumption to monetary policy shocks). Moreover, the model can match empirical findings that are difficult to match with those models (a rapid response of prices to market-specific shocks and a fairly quick response of the price level to aggregate technology shocks). In addition, the model matches all these empirical findings with an endogenous allocation of attention, i.e., with information flows chosen by agents. In principle, one could also match these empirical findings with a model with exogenous dispersed information. One could simply assume the information structure that agents in our model choose in equilibrium (rather than deriving the information structure from an objective and a set of constraints). A natural question to ask is whether the model yields different counterfactuals than the standard New Keynesian model, the sticky information model, and a model with exogenous dispersed information.

The fourth set of predictions concern policy experiments. We find that, due to the endogeneity of the allocation of attention, the outcomes of policy experiments conducted with this model differ markedly from the outcomes of the same policy experiments conducted with other DSGE models. For example, in the model monetary policy is described by a Taylor rule and therefore one can ask what happens when the central bank raises the nominal interest rate more aggressively in response to inflation. In the Calvo model, increasing the coefficient on inflation in the Taylor rule implies that 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. By contrast, in the rational inattention model, there is a non-monotonic relationship between the coefficient on inflation in the Taylor rule and output gap volatility. In our benchmark economy, the standard deviation of the output gap due to monetary policy shocks is essentially constant until a Taylor rule coefficient of 1.75 and then rises. The standard deviation of the output gap due to aggregate technology shocks first rises, peaking at a Taylor rule coefficient of 1.75, and then falls. 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.

Other experiments also yield markedly different outcomes than in other DSGE models. For example, consider increasing the degree of strategic complementarity in price setting. There is a large literature arguing that increasing the degree of strategic complementarity in price setting increases real effects of monetary policy shocks. For example, Woodford (2003), Chapter 3, makes this point for the Calvo model, Mankiw and Reis (2002) make this point for the sticky information model, and Woodford (2002) makes this point for a model with exogenous dispersed information. A common way to increase the degree of strategic complementarity in price setting is to make a firm's marginal cost curve more upward sloping in own output. See Altig, Christiano, Eichenbaum and

Linde (2005). When we increase the degree of strategic complementarity in price setting by making the firms' marginal cost curve more upward sloping in own output, we find that, for reasonable parameter values, real effects of monetary policy shocks become smaller, not larger. The reason is that in the rational inattention model there is an additional effect. When the marginal cost curve becomes more upward sloping in own output, the cost of a price setting mistake of a given size increases. Therefore, decision-makers in firms decide to pay more attention to the price setting decision, implying that prices respond faster to shocks. This effect reduces real effects of monetary policy shocks.

This paper is related to two strands of literature, the literature on rational inattention and the literature on business cycle models with imperfect information. There are several differences to the existing literature on rational inattention (e.g., Sims (2003, 2006), Luo (2008), Maćkowiak and Wiederholt (2009), Woodford (2009), Van Nieuwerburgh and Veldkamp (2009, 2010), and Mondria (2010)). First, this paper develops a dynamic stochastic general equilibrium model with rational inattention. One could interpret the model in Maćkowiak and Wiederholt (2009) as a DSGE model because the model is dynamic, there are multiple firms, and the price level is endogenous, but in that model the demand side of the economy is simply an exogenous process for nominal spending (i.e., households and the central bank are reduced to an exogenous process for nominal spending).[3] By contrast, the model developed here has firms, households, and a central bank; decision-makers in firms and households decide how to allocate their attention; and the central bank sets the nominal interest rate according to a Taylor rule. In addition, there are more shocks which allows us to study the responses of prices to aggregate technology shocks and monetary policy shocks. Second, this paper studies the consumption-saving decision of a household that faces a variable interest rate. Sims (2003, 2006), Luo (2008), and Tutino (2009) also study consumption-saving decisions under rational inattention, and Reis (2006) studies the consumption-saving decision of a household that has to pay a fixed cost to learn the state of the economy. However, in all those papers the real interest rate is constant. Thus, the point that households have little incentive to track movements in the real interest rate (for low and high values of the coefficient of relative risk aversion) is not in

[3] The only monetary policy experiment that can be conducted in that model is a change in the exogenous process for nominal spending. It is unclear what one can learn from this policy experiment for the actual decision problems faced by central banks.

those papers. Moreover, this point is important because in a large class of models monetary policy acts through the real interest rate and therefore in these models the attention that households devote to the real interest rate should be crucial for the speed of response of the economy to monetary policy changes. Third, Paciello (2010) solves a general equilibrium model with rational inattention analytically. The main differences to the model here are that in his model there is only rational inattention on the side of decision-makers in firms and the model is static in the sense that: (i) all exogenous processes are white noise processes, (ii) the price level instead of inflation appears in the Taylor rule, and (iii) there is no lagged interest rate in the Taylor rule.

Compared to the existing literature on business cycle models with imperfect information (e.g., Lucas (1972), Woodford (2002), Mankiw and Reis (2002, 2007), Angeletos and La'O (2009), and Lorenzoni (2009)), the main difference is that information flows are the outcome of an optimization problem (i.e., information flows follow from an objective and a set of constraints). This has two implications. First, the model provides an explanation for equilibrium information flows. Second, the model predicts how information flows vary with policy, which has an important impact on the outcomes of policy experiments. In addition, the model can match empirical findings that other business cycle models with imperfect information have difficulties matching or have not addressed. Namely, the model can match the empirical finding that prices respond rapidly to market-specific shocks as well as the empirical finding that the price level responds faster to aggregate technology shocks than to monetary policy shocks. The sticky information models in Mankiw and Reis (2002, 2007) have difficulties matching these findings. The other papers cited above developing business cycle models with an exogenous information structure do not address these findings.

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 presents the analytical solution of the model under perfect information. Section 7 states the attention problem of the decision-maker in a firm, and presents numerical solutions of the model under rational inattention by decision-makers in firms and perfect information on the side of households. Section 8 states the attention problem of a household, and presents numerical solutions of the model under rational inattention by firms and households. Section 9 concludes.

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_{jt} is composite consumption by household j in period t, L_{jt} is labor supply by household j in period t, and C_{ijt} is consumption of good i by household j in period t. The parameter $\gamma > 0$ is the inverse of the intertemporal elasticity of substitution. 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.[4]

The flow budget constraint of household j in period t reads

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

where P_{it} is the price of good i in period t, B_{jt} are holdings of nominal government bonds by household j between period t and period $t+1$, R_t is the nominal gross interest rate on those bond holdings, W_{jt} is the nominal wage rate for labor supplied by household j in period t, τ_w is a wage subsidy paid by the government, (D_t/J) is a pro-rata share of nominal aggregate profits, and (T_t/J) is a pro-rata share of nominal lump-sum taxes. We assume that all households have the same initial

[4] 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.

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. 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_{it} is composite labor input, L_{ijt} is input of type j labor, 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. The parameter $\eta > 1$ is the elasticity of substitution between those types of labor. The parameter $\alpha \in (0, 1]$ is the elasticity of output with respect to composite labor 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.[5]

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 are assumed to 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 maturing 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

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

[5] Dixit and Stiglitz (1977) also assume that there is a finite number of firms and that firms take the price index as given. Moreover, it seems a good description of the U.S. economy that there is a finite number of firms producing consumption goods and that firms take the consumer price index (CPI) as given.

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.[6] We make this assumption to abstract from the level distortions arising from monopolistic competition.

2.4 Shocks

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

When we aggregate decisions by individual firms, the term $\frac{1}{I}\sum_{i=1}^{I} \varepsilon_{it}^I$ appears. This term is a random variable with mean zero and variance $\frac{1}{I}Var\left(\varepsilon_{it}^I\right)$. When we aggregate individual decisions, we neglect this term because the term has mean zero and a variance that can be made arbitrarily small by increasing the number of firms I.

2.5 Notation

In this subsection we introduce convenient notation. Throughout the paper, C_t denotes aggregate composite consumption

$$C_t = \sum_{j=1}^{J} C_{jt}, \tag{12}$$

and L_t denotes aggregate composite labor input

$$L_t = \sum_{i=1}^{I} L_{it}. \tag{13}$$

[6]When households have perfect information then $\tilde{\theta} = \theta$ and $\tau_p = \frac{\theta}{\theta-1} - 1$. 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.

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

$$\hat{P}_{it} = \frac{P_{it}}{P_t}, \tag{14}$$

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

$$\hat{W}_{jt} = \frac{W_{jt}}{W_t}. \tag{15}$$

In addition, $\tilde{W}_{jt}$ denotes the real wage rate for type j labor

$$\tilde{W}_{jt} = \frac{W_{jt}}{P_t}, \tag{16}$$

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

$$\tilde{W}_t = \frac{W_t}{P_t}. \tag{17}$$

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 for expected profits below when we assume that decision-makers in firms choose the allocation of their attention so as to maximize expected 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.[7]

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

$$C_{it} = \vartheta \left(\frac{P_{it}}{P_t}\right)^{-\tilde{\theta}} C_t, \tag{18}$$

[7]The non-stochastic steady state of the economy presented in Section 2 is characterized in Appendix A. The inflation rate in the non-stochastic steady state 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.

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

$$P_t = d\left(P_{1t}, \ldots, P_{It}\right). \tag{19}$$

The price elasticity of demand $\tilde{\theta} > 1$ is an undetermined coefficient and the constant ϑ equals

$$\vartheta = \hat{P}_i^{-\left(\theta - \tilde{\theta}\right)}, \tag{20}$$

where $\hat{P}_i$ is the relative price of good i in the non-stochastic steady state. In Sections 6-8 when we solve the model for alternative assumptions about information flows, we always verify that this guess concerning the demand function is correct.[8]

Second, we substitute the demand function (18), the production function (4)-(5) and $Y_{it} = C_{it}$ into the expression for profits (6) to obtain the profit function. This yields

$$\left(1 + \tau_p\right) P_{it} \vartheta \left(\frac{P_{it}}{P_t}\right)^{-\tilde{\theta}} C_t - \left[\frac{\vartheta \left(\frac{P_{it}}{P_t}\right)^{-\tilde{\theta}} C_t}{e^{a_t} e^{a_{it}}}\right]^{\frac{1}{\alpha}} \left[\sum_{j=1}^{J-1} W_{jt} \hat{L}_{ijt} + W_{Jt} \left(1 - \sum_{j=1}^{J-1} \hat{L}_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\right]. \tag{21}$$

Profit of firm i in period t equals revenue minus cost. Cost equals the wage bill. The wage bill equals the product of the composite labor input and the wage bill per unit of composite labor input. The wage bill per unit of composite labor input depends on the wages of all types of labor and the labor mix. Profit of firm i in period t depends on the price set by the decision-maker in the firm, P_{it}, the labor mix chosen by the decision-maker in the firm, $\left(\hat{L}_{i1t}, \ldots, \hat{L}_{i(J-1)t}\right)$, and 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 decision-makers in firms in period -1 values nominal profit in period t using the following stochastic discount factor

$$Q_{-1,t} = \beta^t \Lambda\left(C_{1t}, \ldots, C_{Jt}\right) \frac{1}{P_t}, \tag{22}$$

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

where P_t is the price index that appears in the demand function (18) and Λ is some twice continuously differentiable function with the property that the value of this function at the non-stochastic steady state equals the marginal utility of consumption in the non-stochastic steady state[9]

$$\Lambda\left(C_1,\ldots,C_J\right)=C_j^{-\gamma}. \tag{23}$$

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

$$E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t F\left(\hat{P}_{it},\hat{L}_{i1t},\ldots,\hat{L}_{i(J-1)t},a_t,a_{it},C_{1t},\ldots,C_{Jt},\tilde{W}_{1t},\ldots,\tilde{W}_{Jt}\right)\right], \tag{24}$$

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

$$\begin{aligned}
& F\left(\hat{P}_{it},\hat{L}_{i1t},\ldots,\hat{L}_{i(J-1)t},a_t,a_{it},C_{1t},\ldots,C_{Jt},\tilde{W}_{1t},\ldots,\tilde{W}_{Jt}\right)\\
= \; & \Lambda\left(C_{1t},\ldots,C_{Jt}\right)\left(1+\tau_p\right)\vartheta\hat{P}_{it}^{1-\tilde{\theta}}\left(\sum_{j=1}^{J}C_{jt}\right)\\
& -\Lambda\left(C_{1t},\ldots,C_{Jt}\right)\left[\frac{\vartheta\hat{P}_{it}^{-\tilde{\theta}}\left(\sum_{j=1}^{J}C_{jt}\right)}{e^{a_t}e^{a_{it}}}\right]^{\frac{1}{\alpha}}\left[\sum_{j=1}^{J-1}\tilde{W}_{jt}\hat{L}_{ijt}+\tilde{W}_{Jt}\left(1-\sum_{j=1}^{J-1}\hat{L}_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\right].
\end{aligned} \tag{25}$$

Fourth, we compute a log-quadratic approximation to the real profit function around the non-stochastic steady state. In the following, variables without time subscript denote values in the non-stochastic steady state and small variables denote log-deviations from the non-stochastic steady state. 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 (20) as well as the steady state relationships (114), (115), (117), $Y_i=L_i^{\alpha}$ and $Y_i=\hat{P}_i^{-\theta}C$ yields the following real

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

profit function

$$\begin{aligned}
& f\left(\hat{p}_{it}, \hat{l}_{i1t}, \ldots, \hat{l}_{i(J-1)t}, a_t, a_{it}, c_{1t}, \ldots, c_{Jt}, \tilde{w}_{1t}, \ldots, \tilde{w}_{Jt}\right) \\
= \quad & \Lambda\left(C_1 e^{c_{1t}}, \ldots, C_J e^{c_{Jt}}\right) \frac{\tilde{\theta}}{\tilde{\theta}-1} \frac{1}{\alpha} \tilde{W} L_i \frac{1}{J} \sum_{j=1}^{J} e^{\left(1-\tilde{\theta}\right)\hat{p}_{it}+c_{jt}} \\
& -\Lambda\left(C_1 e^{c_{1t}}, \ldots, C_J e^{c_{Jt}}\right) \tilde{W} L_i e^{-\frac{\tilde{\theta}}{\alpha}\hat{p}_{it}-\frac{1}{\alpha}\left(a_t+a_{it}\right)} \left(\frac{1}{J}\sum_{j=1}^{J} e^{c_{jt}}\right)^{\frac{1}{\alpha}} \\
& \frac{1}{J}\left[\sum_{j=1}^{J-1} e^{\tilde{w}_{jt}+\hat{l}_{ijt}} + e^{\tilde{w}_{Jt}}\left(J-\sum_{j=1}^{J-1} e^{\frac{\eta-1}{\eta}\hat{l}_{ijt}}\right)^{\frac{\eta}{\eta-1}}\right]. \qquad (26)
\end{aligned}$$

A second-order Taylor approximation to the real profit function f yields the following proposition. This proposition gives the profit-maximizing decisions and the loss in profit in the case of suboptimal decisions after the log-quadratic approximation to the real profit function.

Proposition 1 *(Expected discounted sum of profits) Let f denote the real profit function given by equation (26). 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 in firm i in period -1. Let x_t, z_t and v_t denote the following vectors*

$$x_t' = \begin{pmatrix} \hat{p}_{it} & \hat{l}_{i1t} & \cdots & \hat{l}_{i(J-1)t} \end{pmatrix}, \qquad (27)$$

$$z_t' = \begin{pmatrix} a_t & a_{it} & c_{1t} & \cdots & c_{Jt} & \tilde{w}_{1t} & \cdots & \tilde{w}_{Jt} \end{pmatrix}, \qquad (28)$$

$$v_t' = \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}, \qquad (29)$$

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

$$E_{i,-1}\left|v_{m,t}v_{n,t}\right| < \delta^t A. \qquad (30)$$

Then the expected discounted sum of profit losses in the case of suboptimal decisions equals

$$E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t \tilde{f}\left(x_t, z_t\right)\right] - E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t \tilde{f}\left(x_t^*, z_t\right)\right] = \sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}\left(x_t - x_t^*\right)' H \left(x_t - x_t^*\right)\right], \qquad (31)$$

where the matrix H is given by

$$H = -C_j^{-\gamma} \tilde{W} L_i \begin{bmatrix} \frac{\tilde{\theta}}{\alpha}\left(1+\frac{1-\alpha}{\alpha}\tilde{\theta}\right) & 0 & \cdots & \cdots & 0 \\ 0 & \frac{2}{\eta J} & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J} \\ \vdots & \frac{1}{\eta J} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1}{\eta J} \\ 0 & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J} & \frac{2}{\eta J} \end{bmatrix}, \tag{32}$$

and the vector x_t^* is given by

$$\hat{p}_{it}^* = \frac{\frac{1-\alpha}{\alpha}}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\left(\frac{1}{J}\sum_{j=1}^{J} c_{jt}\right) + \frac{1}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\left(\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}\right) - \frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\left(a_t + a_{it}\right), \tag{33}$$

and

$$\hat{l}_{ijt}^* = -\eta\left(\tilde{w}_{jt} - \frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}\right). \tag{34}$$

Proof. See Appendix B in Maćkowiak and Wiederholt (2010). ■

After the log-quadratic approximation to the real profit function, the profit-maximizing price in period t is given by equation (33), the profit-maximizing labor mix in period t is given by equation (34), and the loss in profit in period t in the case of a suboptimal decision vector is given by the quadratic form in expression (31). The upper-left element of the matrix H determines the loss in profit in the case of a suboptimal price. The lower-right block of the matrix H determines the loss in profit in the case of a suboptimal factor mix. The loss in profit 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, $(1/\alpha)$, while the loss in profit 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, whereas the off-diagonal elements of H determine how a deviation in one variable affects the loss in profit due to a deviation in another variable. Finally, condition (30) 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 (33)-(34) does not depend at all on the function Λ appearing in the stochastic discount factor (22) because the profit-maximizing price and factor mix are the solution to a static maximization problem. Furthermore, the expected discounted sum of profit losses (31) depends only on the value of the function Λ at 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 profit losses in the case of suboptimal decisions for the economy presented in Section 2 when the demand function is given by equation (18) and the stochastic discount factor is given by equation (22). From this expression one can already see to some extent how the 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 loss in profit in the case of a deviation of the price from the profit-maximizing price (i.e., the attention devoted to the price setting decision will depend on the upper-left element of the matrix H). Moreover, for the decision-maker in a firm it is particularly important to track those changes in the environment that in expectation cause most of the fluctuations in the profit-maximizing decisions. As one can see from equations (33)-(34), which changes in the environment in expectation cause most of the fluctuations in the profit-maximizing decisions depends on the behavior of other agents in the economy, the calibration of the exogenous processes, and the technology parameters α and η. Namely, the price setting behavior of other firms and the consumption and wage setting behavior of households will affect the optimal allocation of attention by the decision-maker in a firm.

4 Derivation of the households' objective

In this section we derive a log-quadratic approximation to the expected discounted sum of period utility. We use this expression for expected utility below when we assume that households choose the allocation of attention so as to maximize expected 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

$$L_{jt} = \zeta \left(\frac{W_{jt}}{W_t} \right)^{-\tilde{\eta}} L_t, \tag{35}$$

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

$$W_t = h \left(W_{1t}, \ldots, W_{Jt} \right). \tag{36}$$

The wage elasticity of labor demand $\tilde{\eta} > 1$ is an undetermined coefficient and the constant ζ equals

$$\zeta = \hat{W}_j^{-(\eta - \tilde{\eta})}. \tag{37}$$

In Sections 6-8 when we solve the model for alternative assumptions about information flows, we always verify that this guess concerning the labor demand function is correct.[10]

Second, we substitute the consumption aggregator (2), the flow budget constraint (3) and the labor demand function (35) into the period utility function (1) to obtain a period utility function that incorporates those constraints. Rearranging the flow budget constraint (3) yields

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

where $\hat{C}_{ijt} = (C_{ijt}/C_{jt})$ is relative consumption of good i and the denominator on the right-hand side is consumption expenditure per unit of composite consumption. Dividing the numerator and the denominator on the right-hand side of the last equation by some price index P_t yields

$$C_{jt} = \frac{\frac{R_{t-1}}{\Pi_t} \tilde{B}_{jt-1} - \tilde{B}_{jt} + (1 + \tau_w) \tilde{W}_{jt} L_{jt} + \frac{\tilde{D}_t}{J} - \frac{\tilde{T}_t}{J}}{\sum_{i=1}^{I} \hat{P}_{it} \hat{C}_{ijt}}, \tag{38}$$

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

$$1 = \sum_{i=1}^{I} \hat{C}_{ijt}^{\frac{\theta-1}{\theta}}. \tag{39}$$

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

Substituting the labor demand function (35), the flow budget constraint (38) and the consumption aggregator (39) into the period utility function (1) yields a period utility function that incorporates those constraints:

$$\frac{1}{1-\gamma}\left(\frac{\frac{R_{t-1}}{\Pi_t}\tilde{B}_{jt-1}-\tilde{B}_{jt}+(1+\tau_w)\tilde{W}_{jt}\zeta\left(\frac{\tilde{W}_{jt}}{\tilde{W}_t}\right)^{-\tilde{\eta}}L_t+\frac{\tilde{D}_t}{J}-\frac{\tilde{T}_t}{J}}{\sum_{i=1}^{I-1}\hat{P}_{it}\hat{C}_{ijt}+\hat{P}_{It}\left(1-\sum_{i=1}^{I-1}\hat{C}_{ijt}^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}}\right)^{1-\gamma}$$
$$-\frac{1}{1-\gamma}-\frac{\varphi}{1+\psi}\left[\zeta\left(\frac{\tilde{W}_{jt}}{\tilde{W}_t}\right)^{-\tilde{\eta}}L_t\right]^{1+\psi}. \tag{40}$$

Third, we compute a log-quadratic approximation to the expected discounted sum of period utility around the non-stochastic steady state. Expressing the period utility function (40) in terms of log-deviations from the non-stochastic steady state and using equations (11) and (37) as well as the steady state relationships (111)-(113), (116) and $L_j=\hat{W}_j^{-\eta}L$ yields the following period utility function

$$\frac{C_j^{1-\gamma}}{1-\gamma}\left(\frac{\frac{\omega_B}{\beta}e^{r_{t-1}-\pi_t+\tilde{b}_{jt-1}}-\omega_B e^{\tilde{b}_{jt}}+\frac{\tilde{\eta}}{\tilde{\eta}-1}\omega_W e^{(1-\tilde{\eta})\tilde{w}_{jt}+\tilde{\eta}\tilde{w}_t+l_t}+\omega_D e^{\tilde{d}_t}-\omega_T e^{\tilde{t}_t}}{\frac{1}{I}\sum_{i=1}^{I-1}e^{\hat{p}_{it}+\hat{c}_{ijt}}+\frac{1}{I}e^{\hat{p}_{It}}\left(I-\sum_{i=1}^{I-1}e^{\frac{\theta-1}{\theta}\hat{c}_{ijt}}\right)^{\frac{\theta}{\theta-1}}}\right)^{1-\gamma}$$
$$-\frac{1}{1-\gamma}-\frac{C_j^{1-\gamma}}{1+\psi}\omega_W e^{-\tilde{\eta}(1+\psi)(\tilde{w}_{jt}-\tilde{w}_t)+(1+\psi)l_t}, \tag{41}$$

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

$$\begin{pmatrix}\omega_B & \omega_W & \omega_D & \omega_T\end{pmatrix}=\begin{pmatrix}\frac{\tilde{B}_j}{C_j} & \frac{\tilde{W}_j L_j}{C_j} & \frac{\frac{\tilde{D}}{J}}{C_j} & \frac{\frac{\tilde{T}}{J}}{C_j}\end{pmatrix}. \tag{42}$$

The next proposition gives the utility-maximizing decisions and the loss in utility in the case of suboptimal decisions after 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 (41) 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

$$\begin{aligned} x_t' &= \begin{pmatrix} \tilde{b}_{jt} & \tilde{w}_{jt} & \hat{c}_{1jt} & \cdots & \hat{c}_{I-1jt} \end{pmatrix}, && (43) \\ z_t' &= \begin{pmatrix} r_{t-1} & \pi_t & \tilde{w}_t & l_t & \tilde{d}_t & \tilde{t}_t & \hat{p}_{1t} & \cdots & \hat{p}_{It} \end{pmatrix}, && (44) \\ v_t' &= \begin{pmatrix} x_t' & z_t' & 1 \end{pmatrix}, && (45) \end{aligned}$$

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

$$E_{j,-1}\left[\tilde{b}_{j,-1}^2\right] < \infty, \tag{46}$$

and for all n,

$$E_{j,-1}\left|\tilde{b}_{j,-1} v_{n,0}\right| < \infty. \tag{47}$$

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

$$E_{j,-1}\left|v_{m,t} v_{n,t+\tau}\right| < \delta^t A. \tag{48}$$

Then the expected discounted sum of utility losses in the case of suboptimal decisions equals

$$\begin{aligned} & E_{j,-1}\left[\tilde{g}\left(\tilde{b}_{j,-1}, x_0, z_0, x_1, z_1, \ldots\right)\right] - E_{j,-1}\left[\tilde{g}\left(\tilde{b}_{j,-1}, x_0^*, z_0, x_1^*, z_1, \ldots\right)\right] \\ = & \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2}\left(x_t - x_t^*\right)' H_0 \left(x_t - x_t^*\right) + \left(x_t - x_t^*\right)' H_1 \left(x_{t+1} - x_{t+1}^*\right)\right]. \end{aligned} \tag{49}$$

Here the matrix H_0 equals

$$H_0 = -C_j^{1-\gamma} \begin{bmatrix} \gamma\omega_B^2\left(1+\frac{1}{\beta}\right) & \gamma\omega_B\tilde{\eta}\omega_W & 0 & \cdots & 0 \\ \gamma\omega_B\tilde{\eta}\omega_W & \tilde{\eta}\omega_W\left(\gamma\tilde{\eta}\omega_W + 1 + \psi\tilde{\eta}\right) & 0 & \cdots & 0 \\ 0 & 0 & \frac{2}{\theta I} & \cdots & \frac{1}{\theta I} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \frac{1}{\theta I} & \cdots & \frac{2}{\theta I} \end{bmatrix}, \tag{50}$$

the matrix H_1 equals

$$H_1 = C_j^{1-\gamma} \begin{bmatrix} \gamma\omega_B^2 & \gamma\omega_B\tilde{\eta}\omega_W & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{51}$$

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 each period $t \geq 0$, the vector x_t^* satisfies

$$c_{jt}^* = E_t\left[-\frac{1}{\gamma}\left(r_t - \pi_{t+1} - \frac{1}{I}\sum_{i=1}^{I}\left(\hat{p}_{it+1} - \hat{p}_{it}\right)\right) + c_{jt+1}^*\right], \tag{52}$$

$$\tilde{w}_{jt}^* = \frac{\gamma}{1+\tilde{\eta}\psi}c_{jt}^* + \frac{\psi}{1+\tilde{\eta}\psi}\left(\tilde{\eta}\tilde{w}_t + l_t\right) + \frac{1}{1+\tilde{\eta}\psi}\left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \tag{53}$$

$$\hat{c}_{ijt}^* = -\theta\left(\hat{p}_{it} - \frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \tag{54}$$

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

$$c_{jt}^* = \frac{\omega_B}{\beta}\left(r_{t-1} - \pi_t + \tilde{b}_{jt-1}^*\right) - \omega_B\tilde{b}_{jt}^* + \frac{\tilde{\eta}}{\tilde{\eta}-1}\omega_W\left[\left(1-\tilde{\eta}\right)\tilde{w}_{jt}^* + \tilde{\eta}\tilde{w}_t + l_t\right] + \omega_D\tilde{d}_t - \omega_T\tilde{t}_t - \left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \tag{55}$$

and E_t denotes the expectation operator conditioned on the entire history of the economy up to and including period t, and (iii) the vector v_t with $x_t = x_t^*$ satisfies conditions (46)-(48).

Proof. See Appendix C in Maćkowiak and Wiederholt (2010). ∎

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 (46)-(48) can be ranked using equation (49). Equations (52)-(55) characterize the decisions that the household would take if the household had perfect information in each period $t \geq 0$. After the log-quadratic approximation to the expected discounted sum of period utility, the optimal decisions under perfect information are given by the usual log-linear first-order conditions. Furthermore, equation (49) gives the loss in expected utility in the case of deviations from the optimal decisions under perfect information. The upper-left blocks of the matrices H_0 and H_1 determine the loss in expected utility in the case of suboptimal real bond holdings and suboptimal real wage rates. According to the (1,1) element of the matrix H_0, a single deviation of real bond holdings from optimal real bond holdings causes a larger utility loss the larger γ, ω_B, and $(1/\beta) = (R/\Pi)$. According to the (2,2) 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, $\tilde{\eta}$, and ω_W. In addition, the off-diagonal elements of H_0 show that a wage deviation in period t affects the utility cost of a bond deviation in period t. Moreover, the first row of H_1 shows that a bond deviation in period t affects both the utility cost of a bond deviation in period $t+1$ and the utility cost of a wage deviation in period $t+1$. The lower-right block of the matrix H_0 determines the loss in expected utility in the case of a suboptimal consumption mix. The loss in expected 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 (46)-(48) 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 utility losses in the case of suboptimal decisions for the economy presented in Section 2 when the labor demand function is given by equation (35). This expression is important because inattention leads to deviations from the optimal decisions under perfect information. To choose the optimal allocation of attention, the household has to compare the cost in terms of expected utility of different types of deviations from the optimal decisions under perfect information. From Proposition 2 one can already see to some extent how parameters affect the optimal allocation of attention by a household. For example, consider the role of γ. Increasing γ raises the utility loss in the case of a given deviation of real bond holdings from optimal real bond holdings. At the same time, increasing γ lowers the response of optimal real bond holdings to the real interest rate. The relative strength of these two effects determines 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 (12) and for aggregate composite labor input (13) yields

$$y_t = \frac{1}{I}\sum_{i=1}^{I}\left(\hat{p}_{it} + y_{it}\right), \tag{56}$$

$$c_t = \frac{1}{J}\sum_{j=1}^{J} c_{jt}, \tag{57}$$

and

$$l_t = \frac{1}{I}\sum_{i=1}^{I} l_{it}. \tag{58}$$

Log-linearizing the equations for the price index (19) and for the wage index (36) yields

$$0 = \sum_{i=1}^{I} \hat{p}_{it}, \tag{59}$$

and

$$0 = \sum_{j=1}^{J} \hat{w}_{jt}. \tag{60}$$

The last two equations can also be stated as

$$p_t = \frac{1}{I}\sum_{i=1}^{I} p_{it}, \tag{61}$$

and

$$w_t = \frac{1}{J}\sum_{j=1}^{J} w_{jt}. \tag{62}$$

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,

$$c_{it} = \frac{1}{J}\sum_{j=1}^{J} c_{ijt}, \tag{63}$$

and

$$l_{jt} = \frac{1}{I}\sum_{i=1}^{I} l_{ijt}. \tag{64}$$

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

$$y_{it} = a_t + a_{it} + \alpha l_{it}, \tag{65}$$

and

$$r_t = \rho_R r_{t-1} + (1-\rho_R)\left(\phi_\pi \pi_t + \phi_y y_t\right) + \varepsilon_t^R. \tag{66}$$

6 Perfect information

Next 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 labor mix; households choose the utility-maximizing consumption vector and wage rate; the government sets the nominal interest rate according to the Taylor 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 (33)-(34), (52)-(55), (56)-(66) 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*

$$y_t = c_t = \frac{1+\psi}{1-\alpha+\alpha\gamma+\psi} a_t, \tag{67}$$

$$l_t = \frac{1-\gamma}{1-\alpha+\alpha\gamma+\psi} a_t, \tag{68}$$

$$\tilde{w}_t = \frac{\gamma+\psi}{1-\alpha+\alpha\gamma+\psi} a_t, \tag{69}$$

$$r_t - E_t\left[\pi_{t+1}\right] = \gamma \frac{1+\psi}{1-\alpha+\alpha\gamma+\psi} E_t\left[a_{t+1} - a_t\right], \tag{70}$$

and

$$\hat{c}_{ijt} = -\theta \hat{p}_{it}, \tag{71}$$

$$\hat{p}_{it} = -\frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\theta} a_{it}, \tag{72}$$

$$\hat{l}_{ijt} = -\eta \hat{w}_{jt}, \tag{73}$$

$$\hat{w}_{jt} = 0. \tag{74}$$

Proof. See Appendix D in Maćkowiak and Wiederholt (2010). ■

Under perfect information aggregate output, aggregate composite consumption, aggregate composite labor input, the real wage index, and the real interest rate are determined by aggregate technology. Relative consumption of good i by household j is determined by firm-specific productivity, and firm i's relative input of type j labor is constant. Importantly, 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 Taylor rule (66) and the real interest rate (70). 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.[11]

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.

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

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

$$\max_{\kappa,B_1(L),B_2(L),B_3(L),C_1(L),C_2(L),C_3(L),\tilde{\eta},\chi} \left\{ \sum_{t=0}^{\infty} \beta^t E_{i,-1} \left[\frac{1}{2} \left(x_t - x_t^*\right)' H \left(x_t - x_t^*\right) \right] - \frac{\mu}{1-\beta}\kappa \right\}, \tag{75}$$

where

$$x_t - x_t^* = \begin{pmatrix} p_{it} \\ \hat{l}_{i1t} \\ \vdots \\ \hat{l}_{i(J-1)t} \end{pmatrix} - \begin{pmatrix} p_{it}^* \\ \hat{l}_{i1t}^* \\ \vdots \\ \hat{l}_{i(J-1)t}^* \end{pmatrix}, \tag{76}$$

subject to the equations characterizing the profit-maximizing decisions

$$p_{it}^* = \underbrace{A_1(L)\,\varepsilon_t^A}_{p_{it}^{A*}} + \underbrace{A_2(L)\,\varepsilon_t^R}_{p_{it}^{R*}} + \underbrace{A_3(L)\,\varepsilon_{it}^I}_{p_{it}^{I*}} \tag{77}$$

$$\hat{l}_{ijt}^* = -\eta \hat{w}_{jt}, \tag{78}$$

the equations characterizing the actual decisions

$$p_{it} = \underbrace{B_1(L)\,\varepsilon_t^A + C_1(L)\,\nu_{it}^A}_{p_{it}^A} + \underbrace{B_2(L)\,\varepsilon_t^R + C_2(L)\,\nu_{it}^R}_{p_{it}^R} + \underbrace{B_3(L)\,\varepsilon_{it}^I + C_3(L)\,\nu_{it}^I}_{p_{it}^I} \tag{79}$$

$$\hat{l}_{ijt} = -\tilde{\eta}\left(\hat{w}_{jt} + \frac{Var\left(\hat{w}_{jt}\right)}{\chi}\nu_{ijt}^L\right), \tag{80}$$

and the constraint on information flow

$$\mathcal{I}\left(\left\{p_{it}^{A*}, p_{it}^{R*}, p_{it}^{I*}, \hat{l}_{i1t}^*, \ldots, \hat{l}_{i(J-1)t}^*\right\}; \left\{p_{it}^A, p_{it}^R, p_{it}^I, \hat{l}_{i1t}, \ldots, \hat{l}_{i(J-1)t}\right\}\right) \le \kappa. \tag{81}$$

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 in the actual decisions are assumed to follow Gaussian white noise processes with unit variance 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 (75) denotes the expectation operator conditioned on the information of the decision-maker in firm i in period -1. We assume that $E_{i,-1}$ is the unconditional expectation operator.

The objective (75) 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. After the log-quadratic approximation to the real profit function, the expected discounted sum of profit losses in the case of suboptimal decisions is given by equation (31). See Proposition 1.[12] The variable $\kappa \geq 0$ is the information flow (attention) devoted to the price setting decision and the factor mix decision. The parameter $\mu \geq 0$ is the per-period marginal cost of information flow. We interpret this cost as an opportunity cost (i.e., devoting more of the scarce resource attention to the price setting decision or the factor mix decision requires paying less of the scarce resource attention to some other decision of the firm that we do not model).

Equations (77)-(78) characterize the profit-maximizing decisions. After the log-quadratic approximation to the real profit function, the profit-maximizing price is given by equation (33) and the profit-maximizing factor mix is given by equation (34). We guess that the profit-maximizing price (33) has the representation (77) after using $p_{it} = \hat{p}_{it} + p_t$ and equations (57) and (62) 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. Rewriting the equation for the profit-maximizing factor mix (34) using equations (62) and $\hat{w}_{jt} = \tilde{w}_{jt} - \tilde{w}_t$ yields equation (78).

Equations (79)-(80) characterize the actual decisions. Consider first equation (79). 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 profit losses due to deviations of the price from the profit-maximizing price but requires a larger information flow. Next, consider equation (80). 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 profit losses 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.[13]

[12] In equation (76), we use the fact that $\hat{p}_{it} - \hat{p}^*_{it} = p_{it} - p^*_{it}$.

[13] We put more structure on the factor mix decision than on the price setting decision. In particular, in equation (80) we express the factor mix as a function of relative wage rates rather than of fundamental shocks. We do this because from equation (80) we derive the labor demand function and a labor demand function specifies labor demand

The constraint on information flow (81) 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 as reduction in uncertainty, where uncertainty is measured by entropy. Entropy is simply a measure of uncertainty. The entropy of a normally distributed random vector $X = (X_1, \ldots, X_N)$ equals

$$H(X) = \frac{1}{2}\log_2\left[(2\pi e)^N \det\Omega_X\right],$$

where $\det\Omega_X$ is the determinant of the covariance matrix of X. Conditional entropy is a measure of conditional uncertainty. If the random vectors $X = (X_1, \ldots, X_N)$ and $Y = (Y_1, \ldots, Y_N)$ have a multivariate normal distribution, the conditional entropy of X given knowledge of Y equals

$$H(X|Y) = \frac{1}{2}\log_2\left[(2\pi e)^N \det\Omega_{X|Y}\right],$$

where $\Omega_{X|Y}$ is the conditional covariance matrix of X given Y. Equipped with measures of uncertainty and conditional uncertainty, one can quantify the information that the random vector Y contains about the random vector X as reduction in uncertainty, $H(X) - H(X|Y)$. The operator $\mathcal{I}$ in the information flow constraint (81) is defined as

$$\mathcal{I}(\{X_t\};\{Y_t\}) = \lim_{T\to\infty}\frac{1}{T}\left[H(X_0, \ldots, X_{T-1}) - H(X_0, \ldots, X_{T-1}|Y_0, \ldots, Y_{T-1})\right], \tag{82}$$

where $\{X_t\}_{t=0}^{\infty}$ and $\{Y_t\}_{t=0}^{\infty}$ are stochastic processes. In words, the operator $\mathcal{I}$ quantifies 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 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[14]

$$\mathcal{I}(\{X_t\};\{Y_t\}) = \lim_{T\to\infty}\frac{1}{T}\left[\frac{1}{2}\log_2\left(\frac{\det\Omega_X}{\det\Omega_{X|Y}}\right)\right]. \tag{83}$$

on and off the equilibrium path. By expressing the labor mix as a function of relative wage rates rather than of fundamental shocks, we specify firm i's relative input of type j labor on and off the equilibrium path.

[14]If X_t is a scalar, then Ω_X is the covariance matrix of the vector $(X_0, \ldots, X_{T-1})$. If X_t is a vector, then Ω_X is the covariance matrix of the vector obtained by stacking the vectors $X_0, \ldots, X_{T-1}$.

Finally, if a variable in the information flow constraint (81) is integrated of order one, we replace the original variable by its first difference in the information flow constraint to ensure that entropy is always finite.[15]

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 (75) is quadratic and the profit-maximizing decisions (77)-(78) follow a Gaussian process.[16] 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 attending to aggregate technology, attending to monetary policy, attending to firm-specific productivity, and attending to relative wage rates are independent activities. We relax this assumption in Section 7.5.

Two remarks are in place before we present solutions of the decision problem (75)-(81). When we solve the decision problem (75)-(81) 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.[17] Furthermore, we evaluate the right-hand side of equation (83) 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 (77) and a guess concerning the stochastic process for the relative wage rate in equation (78). Second, we solve the firms' attention problem (75)-(81). Third, we aggregate the individual prices to obtain the aggregate price level

$$p_t = \frac{1}{I}\sum_{i=1}^{I} p_{it}. \tag{84}$$

[15] If a variable in the information flow constraint (81) 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 (81).

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

[17] We set $p = 2$ and $q = 2$, because we found that increasing p or q further failed to change noticeably the solution of the model. When approximating an infinite-order MA process, we allow the process to have a unit root.

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 (52)-(54), equations (56)-(66), $y_{it} = c_{it}$, and the assumption that aggregate technology follows a first-order autoregressive process imply that the following equations have to be satisfied in equilibrium:

$$c_t = E_t \left[-\frac{1}{\gamma} \left(r_t - p_{t+1} + p_t \right) + c_{t+1} \right], \tag{85}$$

$$\tilde{w}_t = \gamma c_t + \psi l_t, \tag{86}$$

$$y_t = c_t, \tag{87}$$

$$y_t = a_t + \alpha l_t, \tag{88}$$

$$a_t = \rho_A a_{t-1} + \varepsilon_t^A, \tag{89}$$

$$r_t = \rho_R r_{t-1} + (1 - \rho_R) \left[\phi_\pi \left(p_t - p_{t-1} \right) + \phi_y y_t \right] + \varepsilon_t^R, \tag{90}$$

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 (33), $p_{it} = \hat{p}_{it} + p_t$ and equations (57) and (62) imply that the profit-maximizing price is given by

$$p_{it}^* = p_t + \frac{\frac{1-\alpha}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\tilde{\theta}} c_t + \frac{1}{1 + \frac{1-\alpha}{\alpha}\tilde{\theta}} \tilde{w}_t - \frac{\frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\tilde{\theta}} \left(a_t + a_{it} \right). \tag{91}$$

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 (54) and equations (57), (59) and (63) imply that the demand function for good i has the form (18)-(20) 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.[18]

[18] 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

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 (80), (58) and (64) imply that the labor demand function for type j labor has the form (35)-(37) 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 (62) 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

Next we report the numerical solution of the model for the following parameter values. One period in the model is one quarter. 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 (88) 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.[19] We then fit equation (89) 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 estimate the monetary policy rule (90) with the quarterly U.S. data on the Federal Funds rate, inflation, and real GDP from 1960 Q1 to 2006 Q4. We obtain $\rho_R = 0.89$, $\phi_\pi = 1.53$, $\phi_y = 0.33$, and a standard deviation of the innovation

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.

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

equal to 0.0021.[20] 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.[21]

To set the parameters of the process for firm-specific productivity, we follow the recent literature that calibrates menu cost models with firm-specific productivity shocks to U.S. micro price data. Nakamura and Steinsson (2008) and Bils, Klenow and Malin (2009) set the autocorrelation of firm-specific productivity in their monthly models equal to 0.66 and 0.7, respectively. We set the autocorrelation of firm-specific productivity in our quarterly model equal to 0.3 because $(0.3)^{1/3}$ equals a number between 0.66 and 0.7. Klenow and Kryvtsov (2008) report that the median absolute size of price changes excluding sale-related price changes in the U.S. equals 9.7 percent. For this reason, we set 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. This choice yields a standard deviation of the innovation to firm-specific productivity equal to 0.18.[22]

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 (75)-(81). We interpret the cost of information flow as an opportunity cost. The idea is that attention devoted to the price

[20] 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.

[21] We investigated the role of all parameters in the model. We report the effects of changes in parameter values in Section 7.4. Note that restricting the sample to the Great Moderation would have yielded a smaller standard deviation of the innovation in the monetary policy rule. In the model this would imply less attention to monetary policy compared with the benchmark economy.

[22] 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. This implies that less attention is allocated to firm-specific productivity.

setting decision and the factor mix decision could have been devoted to other decisions of the firm that we do not model. We set the marginal cost of information flow equal to 0.1 percent of the firm's steady state revenue. We value the cost of information flow in objective (75) using the value of the stochastic discount factor (22) at the non-stochastic steady state. This yields $\mu = (0.001)(1+\tau_p)\hat{P}_i Y_i C_j^{-\gamma}$. This value for μ will imply that, in equilibrium, the expected per-period loss in profit due to deviations of the price from the profit-maximizing price equals 0.15 percent of the firm's steady state revenue: $(0.0015)(1+\tau_p)\hat{P}_i Y_i$. We find this number reasonable.[23]

We first report the optimal allocation of attention at the equilibrium with rational inattention by decision-makers in firms. The decision-maker in a firm who has to set a price decides to pay most attention to firm-specific productivity, quite a bit of attention to aggregate technology, and little attention to monetary policy. More precisely, of the total attention devoted 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. This equilibrium allocation of attention implies that prices respond rapidly to micro-level shocks, fairly quickly to aggregate technology shocks, and slowly to monetary policy shocks, which matches empirical findings that are difficult to match with other DSGE models. Furthermore, for our choice of the marginal cost of information flow, the attention devoted to the price setting decision is sufficiently high so that the price set by a firm tracks the profit-maximizing price very well. In particular, the expected per-period loss in profit due to deviations of the price from the profit-maximizing price equals 0.15 percent of the firm's steady state revenue.[24] As we will point out below, this number is 30 times smaller than in the Calvo model that yields the same responses of the price level and output to monetary policy shocks.

Figures 1 and 2 show impulse responses of the price level, inflation, output, and the nominal interest rate at the equilibrium with rational inattention by decision-makers in firms and perfect

[23] To illustrate this number, consider the following example. Suppose that the firm with a rationally inattentive decision-maker has a profit margin of 15 percent. If the decision-maker of the firm set the profit-maximizing price in each period, the profit margin would increase to 15.15 percent. Hence, 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).

[24] The expected per-period profit loss due to imperfect tracking of firm-specific productivity equals 0.07 percent of the firm's steady state revenue. The expected per-period profit loss due to imperfect tracking of aggregate technology equals 0.05 percent of the firm's steady state revenue. The expected per-period profit loss due to imperfect tracking of monetary policy equals 0.03 percent of the firm's steady state revenue.

information on the side of households (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. A 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. Under rational inattention by decision-makers in firms, 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 of a monetary policy 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 of a monetary policy shock and then converges slowly to zero. By contrast, under perfect information, the price level adjusts fully on impact of a monetary policy shock, there are no real effects, and the nominal interest rate fails to change.

Consider Figure 2. The price level also shows a dampened and delayed response to an aggregate technology shock (compared with the case of perfect information), but 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 that decision-makers in firms decide to pay about three times as much attention to aggregate technology than to monetary policy. Since the response of the price level to an aggregate technology shock is to some extent dampened and delayed, the output gap is negative for a few quarters after a positive technology shock, implying that output shows a hump-shaped impulse response to an aggregate technology shock.[25]

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

Figures 1-3 show that the model matches simultaneously the following empirical findings: (i) the model matches the empirical finding that the price level responds slowly to monetary policy

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

shocks,[26] (ii) the model matches 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, and (iii) the model matches the empirical finding by Boivin, Giannoni and Mihov (2009) and Maćkowiak, Moench and Wiederholt (2009) that prices respond rapidly to disaggregate shocks. The model matches this combination of fast and slow adjustment of prices to shocks with an endogenous allocation of attention. 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 in micro data and the small variance of the innovation in the Taylor rule. For these parameter values, most of the variance of the profit-maximizing price is due to idiosyncratic shocks, a considerable fraction of the variance of the profit-maximizing price is due to aggregate technology shocks, and only a small fraction of the variance of the profit-maximizing price is due to monetary policy shocks. The decision-maker in a firm who has to set a price therefore pays close attention to firm-specific productivity, quite a bit of attention to aggregate technology, and little attention to monetary policy. In addition, there is an amplification effect because the price level appears in the profit-maximizing price (91). If other firms pay little attention to monetary policy, the profit-maximizing price moves less in response to a monetary policy shock, which reduces the incentive for an individual firm to pay attention to monetary policy.

For comparison, we solved the Calvo model with the same preference, technology and monetary policy parameters. The motivation for this comparison is that the Calvo model is the most commonly used model for monetary policy analysis. We set the Calvo parameter so that prices in the Calvo model change every 2.5 quarters on average, because then the impulse responses to a monetary policy shock are essentially identical in the benchmark economy presented above with rational inattention on the side of decision-makers in firms and in the Calvo model with perfect information. See Figure 4. 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. See Figure 5. 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,

[26] 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).

after a positive 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. Thus, after an aggregate technology shock, the rational inattention model is much closer to a frictionless economy than the Calvo model. Moreover, after a firm-specific productivity shock, the rational inattention model behaves essentially like a frictionless economy.

In the benchmark economy and in the Calvo model, firms experience profit losses due to deviations of the price from the profit-maximizing price. In the benchmark economy, the expected loss in profit due to deviations of the price from the profit-maximizing price is about 30 times smaller than in the Calvo model that yields the same impulse responses of the price level and output to a monetary policy shock.[27] The main reason is that, in the benchmark economy, prices respond slowly to monetary policy shocks, but fairly quickly to aggregate technology shocks, and rapidly to micro-level shocks, whereas in the Calvo model prices respond slowly to all those shocks. To generate a slow response of the price level to monetary policy shocks in the Calvo model, one also has to generate a slow response of prices to other shocks in the Calvo model. In addition, in the rational inattention model deviations of the price from the profit-maximizing price are less likely to be extreme than in the Calvo model.

7.4 The effects of changes in parameter values

We now study whether the model yields different counterfactuals than other DSGE models (e.g., the Calvo model, the sticky information model, and a model with exogenous dispersed information). Does it matter whether one uses this model or another DSGE model for policy analysis? We conduct standard experiments like increasing the coefficient on inflation in the Taylor rule and increasing strategic complementarity in price setting. We find that, due to the endogeneity of the allocation of attention, the outcomes of experiments conducted with this model differ markedly from the outcomes of the same experiments conducted with other DSGE models.

For example, let us vary the coefficient on inflation in the Taylor rule. Figure 6 shows the effect of increasing ϕ_π from 1.05 to 1.5 (our benchmark value) and then to 10 on the volatility of

[27] The expected loss in profit due to suboptimal price responses to idiosyncratic conditions is about 40 times smaller than in the Calvo model. The expected loss in profit due to suboptimal price responses to aggregate conditions is about 20 times smaller than in the Calvo model.

the output gap.[28] We report 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 is essentially constant until 1.75 and then rises. The standard deviation of the output gap due to aggregate technology shocks first rises, peaking at 1.75, and then falls. For comparison, as ϕ_π increases 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. Hence, the rational inattention model yields a markedly different answer to the basic policy question of what happens when the central bank responds more aggressively to inflation.

To understand how the value of ϕ_π affects the economy in the two models, note the following. As ϕ_π increases in the Calvo model, 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 more aggressively with the nominal interest rate to inflation, the price level becomes more stable, implying that decision-makers in firms decide to pay 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. For monetary policy shocks the second effect dominates for values of ϕ_π above 1.75, while for aggregate technology shocks the second effect dominates for values of ϕ_π below 1.75.

Second, consider increasing strategic complementarity in price setting. There is a large literature arguing that increasing strategic complementarity in price setting increases real effects of monetary policy shocks. For example, Woodford (2003), Chapter 3, makes this point for the Calvo model, Mankiw and Reis (2002) make this point for the sticky information model, and Woodford (2002) makes this point for a model with exogenous dispersed information. A common way to increase strategic complementarity in pricing is to make a firm's marginal cost curve more upward sloping in own output. See Altig, Christiano, Eichenbaum and Linde (2005). Therefore, we consider the

[28]Here the output gap is defined as the deviation of aggregate output from equilibrium aggregate output under perfect information. Due to the subsidies (10)-(11) the equilibrium aggregate output under perfect information equals the efficient aggregate output.

experiment of increasing the degree of decreasing returns-to-scale, $(1/\alpha)$. When we decrease α from 1 to 2/3 (our benchmark value) and then to 1/2, real effects of monetary policy shocks first increase and then decrease. The reason is that there are two effects. The first effect is the effect emphasized in the literature cited above. In the benchmark economy, a decrease in α lowers the coefficient on consumption in the equation for the profit-maximizing price. Formally, substituting equations (86)-(88) and $\tilde{\theta} = \theta$ into equation (91) yields the following equation for the profit-maximizing price

$$p_{it}^{*} = p_t + \frac{\frac{1-\alpha}{\alpha} + \gamma + \frac{\psi}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\theta} c_t - \frac{\frac{\psi}{\alpha} + \frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\theta} a_t - \frac{\frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\theta} a_{it}. \tag{92}$$

A decrease in α lowers the coefficient on consumption in equation (92) if and only if $\theta(\gamma + \psi) > (1 + \psi)$, which is a parameter restriction that is satisfied in the benchmark economy. In the language of Ball and Romer (1990), a decrease in α raises the degree of real rigidity, and in the language of Woodford (2003), a decrease in α raises the degree of strategic complementarity in price setting. This effect increases real effects of monetary policy shocks. 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. Decision-makers in firms thus decide to pay more attention to the price setting decision, implying that prices respond faster to shocks. This effect reduces real effects of monetary policy shocks. We find that the second effect (more attention) dominates the first effect (lower coefficient on consumption in the equation for the profit-maximizing price) for values of α below 2/3. Hence, for reasonable parameter values, increasing strategic complementarity reduces real effects.

Third, consider increasing the variance of monetary policy shocks. In the rational inattention model, decision-makers in firms decide to pay more attention to monetary policy, implying that prices respond faster to monetary policy shocks and real effects of a monetary policy shock of a given size decrease. By contrast, in the Calvo model and in the sticky information model of Mankiw and Reis (2002), increasing the variance of monetary policy shocks has no effect on the responses of prices and output to a monetary policy shock of a given size. The reallocation of attention in the rational inattention model is important quantitatively. For example, in the benchmark economy, doubling the standard deviation of monetary policy shocks implies that real effects of monetary policy shocks last only 4 quarters instead of 10 quarters.

One could go on and on with more experiments. The point is: the outcomes of experiments are markedly different than in other DSGE models. Moreover, there is a clear intuition for why the outcomes are so different: the allocation of attention varies with the economic environment.

7.5 Extension: Signals

In this subsection 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.

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:

$$\max_{(\kappa,\sigma_1,\sigma_2,\sigma_3,\sigma_4)\in\mathbf{R}^5_+} \left\{ \sum_{t=0}^{\infty} \beta^t E_{i,-1} \left[\frac{1}{2} \left(x_t - x_t^*\right)' H \left(x_t - x_t^*\right) \right] - \frac{\mu}{1-\beta}\kappa \right\}, \tag{93}$$

where

$$x_t - x_t^* = \begin{pmatrix} p_{it} \\ \hat{l}_{i1t} \\ \vdots \\ \hat{l}_{i(J-1)t} \end{pmatrix} - \begin{pmatrix} p_{it}^* \\ \hat{l}_{i1t}^* \\ \vdots \\ \hat{l}_{i(J-1)t}^* \end{pmatrix}, \tag{94}$$

subject to equations (77)-(78) characterizing the profit-maximizing decisions, the following equation characterizing the optimal decision vector in period t given information in period t

$$x_t = E\left[x_t^* | \mathcal{F}_{i0}, s_{i1}, s_{i2}, \ldots, s_{it}\right], \tag{95}$$

the following equation characterizing the signal vector in period t

$$s_{it} = \begin{pmatrix} p_{it}^{A*} \\ p_{it}^{R*} \\ p_{it}^{I*} \\ \hat{w}_{1t} \\ \vdots \\ \hat{w}_{(J-1)t} \end{pmatrix} + \begin{pmatrix} \sigma_1 \nu_{it}^{A} \\ \sigma_2 \nu_{it}^{R} \\ \sigma_3 \nu_{it}^{I} \\ \sigma_4 \nu_{i1t}^{L} \\ \vdots \\ \sigma_4 \nu_{i(J-1)t}^{L} \end{pmatrix}, \tag{96}$$

and the constraint on information flow

$$\mathcal{I}\left(\left\{p_{it}^{A*}, p_{it}^{R*}, p_{it}^{I*}, \hat{l}_{i1t}^{*}, \ldots, \hat{l}_{i(J-1)t}^{*}\right\}; \{s_{it}\}\right) \leq \kappa. \tag{97}$$

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 Gaussian white noise processes with unit variance 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 (75)-(81), $E_{i,-1}$ in objective (93) 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 (93) is the marginal cost of information flow, and the operator $\mathcal{I}$ in constraint (97) is defined by equation (82). We assume that $E_{i,-1}$ is the unconditional expectation operator. Finally, $\mathcal{F}_{i0}$ in equation (95) 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 covariance matrix of x_t^* given information in period t is constant for all $t \geq 0$.

We solve the problem (93)-(97) 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 (93)-(97) to the solution to problem (75)-(81). Consider the left column of 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 (75)-(81). 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 (93)-(97). The point is that 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 (79) and to the noise terms in equation (96) also turn out to be identical. In summary, the decision problem (75)-(81) and the decision problem (93)-(97) yield the same price setting behavior.[29,30]

The signal vector (96) captures the idea that paying attention to aggregate technology, paying 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

[29] We solve problem (93)-(97) 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 (96) with a finite-dimensional state vector. We use the Kalman filter to evaluate objective (93) and constraint (97) 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 covariance matrix of the state vector. Solving the problem (93)-(97) takes about as much time as solving the problem (75)-(81). See Footnote 18. Below we also present solutions of problem (93)-(97) with the signal vector (98) instead of the signal vector (96). Solving that problem turned out to be much more time-consuming. Here we had to evaluate objective (93) and constraint (97) on a grid. Standard nonlinear optimization programs proved unhelpful because numerical inaccuracy in the solution for the conditional covariance matrix of the state vector led to spurious variation in the values of the objective and the constraint.

[30] This is a numerical result. The fact that there exist some signals that yield the same price setting behavior is not surprising. See Section V in Maćkowiak and Wiederholt (2009). What is surprising is that signals with noise that is i.i.d. across time yield the same price setting behavior as the decision problem (75)-(81).

signal vector (96) by the following signal vector[31]

$$s_{it} = \begin{pmatrix} p_t \\ a_t + a_{it} \\ c_{i,t-1} \\ w_{t-1} + l_{i,t-1} \\ \hat{w}_{1t} \\ \vdots \\ \hat{w}_{(J-1)t} \end{pmatrix} + \begin{pmatrix} \sigma_1 \nu_{i1t} \\ \sigma_2 \nu_{i2t} \\ \sigma_3 \nu_{i3t} \\ \sigma_4 \nu_{i4t} \\ \sigma_5 \nu_{i1t}^L \\ \vdots \\ \sigma_5 \nu_{i(J-1)t}^L \end{pmatrix}. \tag{98}$$

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.[32] The variables in the signal vector (98) 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 (93)-(97) with the signal vector (98) instead of the signal vector (96) changes the firm's price setting behavior hardly at all.[33] See the right column of Figure 7. 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 the right column of Figure 7 are very similar to the red lines with crosses in the left column of Figure 7. The reason is that the decision-maker in the firm decides to pay attention to those variables that are mainly driven by firm-specific productivity shocks and aggregate technology shocks.

We studied a large number of variations of the signal vector (98) and obtained similar results.

[31] We maintain the assumption that the noise terms in the signal follow Gaussian white noise processes with unit variance that are: (i) independent of all other stochastic processes in the economy, (ii) firm-specific, and (iii) independent of each other.

[32] 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.

[33] When we replace the signal vector (96) by the signal vector (98), 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.

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 (98) we replaced last period sales and last period wage bill by current period sales and current period wage bill in the signal vector. The 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 (93)-(97) with the signal vector (98). 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 the right column of Figure 7.[34]

8 Rational inattention by firms and households

In this section we study the implications of adding rational inattention by households. We solve the model with rational inattention by decision-makers in firms and households.

In Sections 8.1-8.4, we assume that households have linear disutility of labor and households set real wage rates. We make these two assumptions because they allow us to exhibit in the most transparent way the implications of rational inattention by households for the consumption-saving decision. The reason is the following. When households have linear disutility of labor ($\psi = 0$), the intratemporal optimality condition stating that the real wage rate should equal the marginal rate

[34] 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. More importantly, in their model the noise in the signal is exogenous, whereas in our model the noise in the signal (98) is chosen optimally subject to the constraint on information flow (97). 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.

of substitution between consumption and leisure reduces to $\tilde{w}_{jt} = \gamma c_{jt}$. Thus, when $\psi = 0$ 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. Hence, when $\psi = 0$ and households set real wage rates, households satisfy this intratemporal optimality condition both under perfect information and under rational inattention. This allows us to exhibit in the most transparent way the implications of rational inattention by households for the intertemporal consumption decision. In Section 8.5, we present the solution when households set nominal wage rates.

8.1 The households' attention problem

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

$$\max_{\kappa, B_1(L), B_2(L), C_1(L), C_2(L), \tilde{\theta}, \xi} \left\{ \begin{array}{c} \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[\frac{1}{2} \left(x_t - x_t^* \right)' H_0 \left(x_t - x_t^* \right) + \left(x_t - x_t^* \right)' H_1 \left(x_{t+1} - x_{t+1}^* \right) \right] \\ - \frac{\lambda}{1-\beta} \kappa \end{array} \right\}, \tag{99}$$

where

$$x_t - x_t^* = \begin{pmatrix} \tilde{b}_{jt} \\ \tilde{w}_{jt} \\ \hat{c}_{1jt} \\ \vdots \\ \hat{c}_{I-1jt} \end{pmatrix} - \begin{pmatrix} \tilde{b}_{jt}^* \\ \tilde{w}_{jt}^* \\ \hat{c}_{1jt}^* \\ \vdots \\ \hat{c}_{I-1jt}^* \end{pmatrix}, \tag{100}$$

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

$$\tilde{b}_{jt} - \tilde{b}_{jt}^* = -\sum_{l=0}^{t} \left(\frac{1}{\beta} \right)^l \frac{1}{\omega_B} \left[\left(c_{j,t-l} - c_{j,t-l}^* \right) + \tilde{\eta} \omega_W \left(\tilde{w}_{j,t-l} - \tilde{w}_{j,t-l}^* \right) \right], \tag{101}$$

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

$$c_{jt}^* = \underbrace{A_1(L) \varepsilon_t^A}_{c_{jt}^{A*}} + \underbrace{A_2(L) \varepsilon_t^R}_{c_{jt}^{R*}} \tag{102}$$

$$\tilde{w}_{jt}^* = \gamma c_{jt}^* \tag{103}$$

$$\hat{c}_{ijt}^* = -\theta \hat{p}_{it}, \tag{104}$$

the equations characterizing the household's actual decisions

$$\begin{aligned} c_{jt} &= \underbrace{B_1(L)\,\varepsilon_t^A + C_1(L)\,\nu_{jt}^A}_{c_{jt}^A} + \underbrace{B_2(L)\,\varepsilon_t^R + C_2(L)\,\nu_{jt}^R}_{c_{jt}^R} && (105)\\ \tilde{w}_{jt} &= \gamma c_{jt} && (106)\\ \hat{c}_{ijt} &= -\tilde{\theta}\left(\hat{p}_{it} + \frac{Var(\hat{p}_{it})}{\xi}\nu_{ijt}^I\right), && (107) \end{aligned}$$

and the constraint on information flow

$$\mathcal{I}\left(\left\{c_{jt}^{A*}, c_{jt}^{R*}, \hat{c}_{1jt}^{*}, \ldots, \hat{c}_{I-1jt}^{*}\right\}; \left\{c_{jt}^{A}, c_{jt}^{R}, \hat{c}_{1jt}, \ldots, \hat{c}_{I-1jt}\right\}\right) \leq \kappa. \tag{108}$$

Here $A_1(L)$, $A_2(L)$, $B_1(L)$, $B_2(L)$, $C_1(L)$ and $C_2(L)$ are infinite-order lag polynomials. The noise terms ν_{jt}^A, ν_{jt}^R, and ν_{ijt}^I in the actual decisions are assumed to follow Gaussian white noise processes with unit variance 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 (82). Finally, $E_{j,-1}$ in objective (99) is the expectation operator conditioned on the information of household j in period -1.

The objective (99) states that the household chooses level and 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.[35] The variable $\kappa \geq 0$ is the overall information flow devoted to the intertemporal consumption decision, the intratemporal consumption decision, and the wage setting decision. The parameter $\lambda \geq 0$ is the per-period marginal cost of information flow. We interpret this cost as an opportunity cost. To devote more attention to the questions of how much to consume, which goods to consume, and which wage to set, the household has to devote less attention to some other activity.

[35] Proposition 2 states that, after the log-quadratic approximation to expected lifetime utility and for sequences satisfying conditions (46)-(48), maximizing expected lifetime utility is equivalent to maximizing the expression on the right-hand side of equation (49). When we solve the households' attention problem (99)-(108), we consider only stochastic processes for real bond holdings, the real wage rate, and the consumption mix that satisfy conditions (46)-(48). It is important to note that conditions (46)-(48) do not require that the processes for real bond holdings, the real wage rate, and the consumption mix are stationary. Conditions (46)-(48) do require that second moments increase less than exponentially in t.

Equations (102)-(104) characterize the household's optimal decisions under perfect information (i.e., the decisions that the same household would take if the household had perfect information in each 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 (52)-(55). See Proposition 2. We guess that c_{jt}^{*} given by equation (52) has the representation (102) after substituting in the equilibrium law of motion for r_t and π_t. The guess will be verified. Equations (53) and (54) reduce to equations (103) and (104) after substituting in equation (59) and $\psi = 0$.

Equations (105)-(107) characterize the household's actual decisions. Consider first equation (105). 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 intertemporal consumption decision in each period. The basic trade-off is the following. Choosing a process for composite consumption that tracks more closely optimal composite consumption under perfect information reduces losses in expected utility due to suboptimal intertemporal consumption decisions but requires a larger information flow. Next, consider equation (106). This equation states that in each 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 (106) is that information contained in the household's own current and past consumption decisions is also used in the household's current wage setting decision. More precisely, in Appendix E in Maćkowiak and Wiederholt (2010) we show analytically that if the household can choose the process for the real wage rate $\{\tilde{w}_{jt}\}_{t=0}^{\infty}$ as a time-invariant one-sided linear filter of the process $\left\{c_{jt}^{A}, c_{jt}^{R}\right\}_{t=0}^{\infty}$, then the optimal filter is equation (106) so long as the household has linear disutility of labor ($\psi = 0$). Finally, consider equation (107). 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 again the following. Choosing a process for the consumption mix that tracks more closely the optimal consumption mix under perfect information reduces losses in expected utility due to suboptimal consumption baskets but requires a larger information flow.[36]

[36] We put more structure on the consumption mix decision than on the intertemporal consumption decision and the wage setting decision. In particular, in equation (107) we express the consumption mix as a function of relative

The constraint on information flow (108) states that actual decisions containing more information about the optimal decisions under perfect information require a larger information flow.

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 had under perfect information is then given by equation (101). Equation (101) follows from equation (55) and $\tilde{b}_{j,-1} = \tilde{b}^*_{j,-1}$. Equation (101) is needed because the deviation $\tilde{b}_{jt} - \tilde{b}^*_{jt}$ is an argument of objective (99). When the household consumes more than the household would have consumed under perfect information, bond holdings are lower than they would have been under perfect information. Note that, since equation (55) is the log-linearized flow budget constraint, equation (101) determines log bond holdings. Log bond holdings may be negative, but bond holdings themselves are always strictly positive.

Finally, we have to specify the expectation operator $E_{j,-1}$ in objective (99). 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 made in Section 2 that all households have the same bond holdings in period -1. Second, this assumption implies that all the discounted second moments in objective (99) are finite even when $(x_t - x^*_t)$ has a unit root, and we want to allow for the possibility that $(x_t - x^*_t)$ has a unit root.

When we solve the problem (99)-(108) 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.[37] Furthermore, when we solve the problem (99)-(108) numerically, we evaluate the right-hand side of equation (83) for a very large but finite T.

---

prices rather than of fundamental shocks. We do this because from equation (107) 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 rather than of fundamental shocks, we specify relative consumption of good i by household j on and off the equilibrium path.

[37] We set $p = 2$ and $q = 2$, because we found that increasing p or q further failed to change noticeably the solution of the model. When approximating an infinite-order MA process, we allow the process to have a unit root.

8.2 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, p_{it}^*, and a guess concerning the stochastic process for the utility-maximizing composite consumption, c_{jt}^*. Second, we solve the firms' attention problem (75)-(81) and we solve the households' attention problem (99)-(108). Third, we aggregate the individual prices to obtain the price level. We aggregate across households to obtain aggregate composite consumption, $c_t = \frac{1}{J}\sum_{j=1}^{J} c_{jt}$, and the real wage index, $\tilde{w}_t = \frac{1}{J}\sum_{j=1}^{J} \tilde{w}_{jt}$. Fourth, we compute the law of motion for the nominal interest rate from the monetary policy rule (90) and equation (87); we compute the law of motion for the profit-maximizing price from equation (91); and we compute the law of motion for the utility-maximizing composite consumption from equation (52). 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.[38]

8.3 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. These parameters 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. These five parameters appear in objective (99). The parameters ω_B, ω_W, and $\tilde{\eta}$ also appear in equation (101) 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, since we want to base our calibration of ω_B and ω_W on data for "typical" U.S. households, we compute median nominal net worth, median nominal annual income, and median nominal annual wage income for the households in the 40-60 income percentile of the SCF 2007. These three statistics

[38] One iteration takes about 4 minutes on the machine described in Footnote 18.

equal \$88400, \$47305, and \$41135, 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, income, and wage income) and the household that is the median household according to one variable may be an unusual household according to the other variables. Second, since consumption appears in the denominator of ω_B and ω_W but the SCF has only very limited data on consumption expenditure, we calculate a proxy for consumption expenditure. 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 nominal annual income given above and we deduct 2.5 percent of nominal net worth given above. This proxy for annual consumption expenditure equals \$38782. Third, we divide annual nominal wage income given above by four to obtain quarterly nominal wage income. 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: $\omega_W = (10283.75/9695.5) = 1.06$. 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 $\tilde{\eta} = 4$. With rational inattention on the side of decision-makers in firms and households, decision-makers on the demand side of each market have rational inattention. For this reason, 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.[39]

We set the number of consumption goods to $I = 1000$. The parameter I has no effect on the responses of the household's composite consumption and the household's real wage rate to shocks.

[39] A price elasticity of demand of four is within the range of estimates of the price elasticity of demand in the Industrial Organization literature.

The parameter I only affects the household's choice of $\tilde{\theta}$ and ξ. Put differently, the parameter I only affects the parameter 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 for the marginal cost of information flow will imply that, in equilibrium, the expected per-period loss in utility due to deviations of composite consumption and the real wage rate from the optimal decisions under perfect information equals the utility equivalent of 0.06 percent of the household's steady state consumption. Put differently, to fully compensate the household for the expected discounted sum of utility losses due to deviations of composite consumption and the real wage rate from the optimal decisions under perfect information, it would be sufficient to give the household 1/1700 of the household's steady state consumption in every period. We think these are very small utility losses.

We first solve the household's attention problem (99)-(108) assuming that aggregate variables and relative prices are given by the equilibrium of the benchmark economy presented in Section 7.3. That is, we study the optimal allocation of attention of an individual household when decision-makers in firms have limited attention and all other households have perfect information. Figure 8 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 under rational inattention. The green lines with circles are the impulse responses under perfect information (i.e., the green lines with circles show what the household would do if the household had perfect information). The impulse responses of consumption to shocks under rational inattention are very different from the impulse responses of consumption to shocks under perfect information, despite the fact that for our parameter values the expected per-period loss in utility due to deviations of composite consumption and the real wage rate from the optimal decisions under perfect information is very small. Importantly, the impulse response of consumption to a monetary policy shock is hump-shaped under rational inattention, whereas the impulse response of consumption to a monetary policy shock is monotonic under perfect information. Furthermore, 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, 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 (not reported here) goes to zero.[40] In summary, under rational inattention composite consumption responds very slowly to shocks. If the household had perfect information in each period like all other households, then composite consumption of the household would equal the sum of current and future real interest rates (i.e., the long rate). The fact that the household responds very slowly with composite consumption to shocks reflects the fact that the household decides to track movements in the real interest rate imperfectly.

One might think that the result that a rational inattention household pays little attention to the intertemporal consumption decision is due to the fact that the coefficient of relative risk aversion is low, implying that deviations from the consumption Euler equation are cheap in utility terms. Therefore, we studied what happens when we increase γ by a factor of 10 from our benchmark value of $\gamma = 1$. As γ increases from 1 to 10, the attention devoted to the intertemporal consumption decision increases by 50 percent and the ratio of the actual response to the optimal response of consumption on impact of a monetary policy shock increases from 12 percent to 26 percent. The household devotes more attention to the intertemporal consumption decision and therefore consumption responds faster to a monetary policy shock. However, note that both for $\gamma = 1$ and for $\gamma = 10$, the household devotes little attention to the intertemporal consumption decision and consumption responds slowly to a monetary policy shock. This is because there are two effects working in opposite directions. Increasing γ raises utility losses in the case of deviations of composite consumption from optimal composite consumption under perfect information. See equation (49). This effect raises the attention devoted to the intertemporal consumption decision. On the other hand, increasing γ lowers the coefficient on the real interest rate in the consumption Euler equation, implying that being aware of movements in the real interest rate becomes less important. See equation (52). This effect lowers the attention devoted to the intertemporal consumption decision. For γ between 1 and 10, the first effect dominates, but only slightly.

Next, we present the equilibrium of the model under rational inattention by decision-makers

[40] We also find that the impulse responses of composite consumption and real bond holdings under rational inattention to the noise terms in equation (105) go to zero in the long run. In the version of the model where all households solve the problem (99)-(108), this finding implies that neither the cross-sectional variance of consumption nor the cross-sectional variance of real bond holdings diverges to infinity.

in firms and rational inattention by households. We use the benchmark parameter values. We compute the rational expectations equilibrium using the iterative procedure described in Section 8.2. Figure 9 shows the impulse responses of the price level, inflation, aggregate composite consumption, and the nominal interest rate to a monetary policy shock. How do the impulse responses to a monetary policy shock change when we add rational inattention by households? First, the impulse response of aggregate composite consumption to a monetary policy shock becomes hump-shaped. This is because households choose to pay little attention to movements in the real interest rate and therefore respond slowly with their composite consumption to a monetary policy shock. Second, the impulse response of the price level to a monetary policy shock becomes even more dampened and delayed compared with the case of rational inattention by decision-makers in firms and perfect information by households. The dampened and delayed response of aggregate composite consumption to monetary policy shocks makes decision-makers in firms pay even less attention to monetary policy, implying that the price level responds even more slowly to a monetary policy shock. Households' optimal allocation of attention affects firms' optimal allocation of attention. Third, the two effects described above are counteracted to some extent by the Taylor rule. The dampened response of consumption and inflation on impact of a monetary policy shock implies that the nominal interest rate responds more strongly on impact of a monetary policy shock. This increases to some extent the size of the response of consumption and inflation to a monetary policy shock.

Figure 10 shows the impulse responses of the price level, inflation, aggregate composite consumption, and the nominal interest rate to an aggregate technology shock. How do the impulse responses to an aggregate technology shock change when we add rational inattention by households? The main change is that the impulse response of aggregate composite consumption becomes even more dampened and delayed. This change is important quantitatively. In the case of rational inattention by decision-makers in firms and perfect information by households, the growth rate of aggregate composite consumption conditional on an aggregate technology shock has a serial correlation of 0.38. When we add rational inattention by households, this number more than doubles, to 0.77. Carroll, Slacalek, and Sommer (2008) estimate that the growth rate of aggregate consumption has a serial correlation of about 0.7, on average across countries. Their estimate for the U.S. is 0.83. This means that, once we add rational inattention by households, the model can match the large serial correlation of aggregate consumption growth in the data.

8.4 The effects of changes in parameter values

When we recompute the experiments reported in Section 7.4 with rational inattention by decision-makers in firms and rational inattention by households, we obtain two main findings. The first finding is qualitative. We confirm that the outcomes of experiments conducted with this model differ markedly from the outcomes of the same experiments conducted with other DSGE models. The reason remains that highlighted in Section 7.4: the allocation of attention varies with the economic environment. The second finding is quantitative. The outcome of a particular experiment may change in an important way after one adds rational inattention by households.

For example, recall that in Section 7.4 we point out two effects of an increase in the coefficient on inflation in the Taylor rule on the volatility of the output gap. First, there is the standard effect. As ϕ_π increases, the nominal interest rate mimics more closely the real interest rate at the efficient solution. Second, there is the effect due to the optimal allocation of attention by decision-makers in firms. As ϕ_π increases, decision-makers in firms decide to pay less attention to aggregate conditions. In Section 7.4, we find that in the case of monetary policy shocks the second effect dominates for values of ϕ_π above 1.75. We find that the standard deviation of the output gap due to monetary policy shocks is essentially constant until $\phi_\pi = 1.75$ and then rises. When we add rational inattention by households, a third effect arises. As ϕ_π increases, the amount of attention that households allocate to aggregate conditions first rises and then falls. The amount of attention that decision-makers in firms allocate to aggregate conditions falls monotonically, as before. In equilibrium, the standard deviation of the output gap due to monetary policy shocks first rises, peaking at $\phi_\pi = 1.5$, and then falls.

As another example, recall that in Section 7.4 we find that, as ϕ_π increases, the standard deviation of the output gap due to aggregate technology shocks first rises, peaking at 1.75, and then falls. After adding rational inattention by households, we find that the peak occurs at $\phi_\pi = 3$. The same three effects interact: the standard effect, the effect that the amount of attention that decision-makers in firms allocate to aggregate conditions falls monotonically with ϕ_π, and the effect that the amount of attention that households allocate to aggregate conditions varies non-monotonically with ϕ_π.

These findings suggest that the interaction between decision-makers in firms and households under rational inattention is important for the outcomes of experiments.

8.5 Extension: Households set nominal wage rates

We have also solved the model assuming households set nominal wage rates instead of real wage rates. See Sections 8.3 and 8.4 in Maćkowiak and Wiederholt (2010). The main change is that rational inattention by households now also causes deviations from the households' intratemporal optimality condition stating that the real wage rate should equal the marginal rate of substitution between consumption and leisure. This has two implications. First, since inattention to aggregate conditions now also causes deviations from the households' intratemporal optimality condition, households decide to pay somewhat more attention to aggregate conditions. This effect tends to make the response of aggregate composite consumption to shocks somewhat stronger and faster. On the other hand, since households set nominal wage rates instead of real wage rates and households pay limited attention to aggregate conditions, the response of wage rates to shocks becomes more dampened and delayed. This effect increases real effects of monetary policy shocks and increases the distance between the efficient response and the actual response of output to an aggregate technology shock. We chose to present the results with households setting real wage rates here because we think that this version of the model exhibits in the most transparent way the effects of rational inattention by households on the intertemporal consumption decision.

9 Conclusion

We develop and solve a DSGE model in which decision-makers in firms and households have limited attention and decide how to allocate their attention. We find that impulse responses to aggregate shocks display substantial inertia, despite the fact that profit losses and utility losses due to rational 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. Moreover, our model stands in stark contrast to standard business cycle models when it comes to the mix of fast and slow adjustment of prices to shocks, profit losses due to deviations of the actual price from the profit-maximizing price, and the outcomes of policy experiments.

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 to the model and to estimate the two parameters that govern slow adjustment, the marginal cost of information flow for the decision-maker in a firm and the marginal cost of information flow for a household.

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

$$P_t = \left(\sum_{i=1}^{I} P_{it}^{1-\theta} \right)^{\frac{1}{1-\theta}}, \tag{109}$$

and W_t denotes the following wage index

$$W_t = \left(\sum_{j=1}^{J} W_{jt}^{1-\eta} \right)^{\frac{1}{1-\eta}}. \tag{110}$$

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

$$\frac{R}{\Pi} = \frac{1}{\beta}, \tag{111}$$

$$\frac{C_{ij}}{C_j} = \hat{P}_i^{-\theta}, \tag{112}$$

and

$$\tilde{W}_j = \varphi \left(\hat{W}_j^{-\eta} L \right)^{\psi} C_j^{\gamma}. \tag{113}$$

The firms' first-order conditions read

$$\hat{P}_i = \tilde{W} \frac{1}{\alpha} \left(\hat{P}_i^{-\theta} C \right)^{\frac{1}{\alpha}-1}, \tag{114}$$

and

$$\hat{L}_{ij} = \hat{W}_j^{-\eta}. \tag{115}$$

The firms' price setting equation (114) 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 (109), the consumption aggregator (2) and the definition of aggregate composite labor input (13) that

$$\hat{P}_i^{1-\theta} = \left(\frac{C_{ij}}{C_j}\right)^{\frac{\theta-1}{\theta}} = \frac{L_i}{L} = \frac{1}{I}. \tag{116}$$

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 (12), the definition of the wage index (110) and the labor aggregator (5) that

$$\frac{C_j}{C} = \hat{W}_j^{1-\eta} = \hat{L}_{ij}^{\frac{\eta-1}{\eta}} = \frac{1}{J}. \tag{117}$$

One can show that equations (111)-(117), $Y_i = L_i^{\alpha}$ and $Y_i = \hat{P}_i^{-\theta} C$ imply that all variables appearing in equations (111)-(117) are uniquely determined apart from the nominal interest rate, R, and inflation, Π. For ease of exposition, we select $\Pi = 1$. Equation (111) 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).

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, 319-378.

[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] Carroll, Christopher D., Jirka Slacalek, and Martin Sommer (2008): "International Evidence on Sticky Consumption Growth." NBER Working Paper 13876.

[7] 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.

[8] 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.

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

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

[11] 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.

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

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

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

[15] Lorenzoni, Guido (2009): "A Theory of Demand Shocks." American Economic Review, 99(5), 2050-2084.

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

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

[18] Maćkowiak, Bartosz, and Mirko Wiederholt (2010): "Business Cycle Dynamics under Rational Inattention." CEPR Discussion Paper 7691.

[19] 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.

[20] Mankiw, N. Gregory, and Ricardo Reis (2007): "Sticky Information in General Equilibrium." Journal of the European Economic Association, 5 (2-3), 603-613.

[21] Mondria, Jordi (2010): "Portfolio Choice, Attention Allocation, and Price Comovement." Discussion paper, University of Toronto.

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

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

[24] Reis, Ricardo (2006): "Inattentive Consumers." Journal of Monetary Economics, 53(8), 1761-1800.

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

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

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

[28] 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.

[29] Tutino, Antonella (2009): "The Rigidity of Choice. Lifecycle Savings with Information Processing Limits." Discussion paper, Federal Reserve Board.

[30] 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.

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

[32] Van Nieuwerburgh, Stijn, and Laura Veldkamp (2010): "Information Acquisition and Under-Diversification." Review of Economic Studies, forthcoming.

[33] 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.

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

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

Figure 1: Impulse responses, benchmark economy

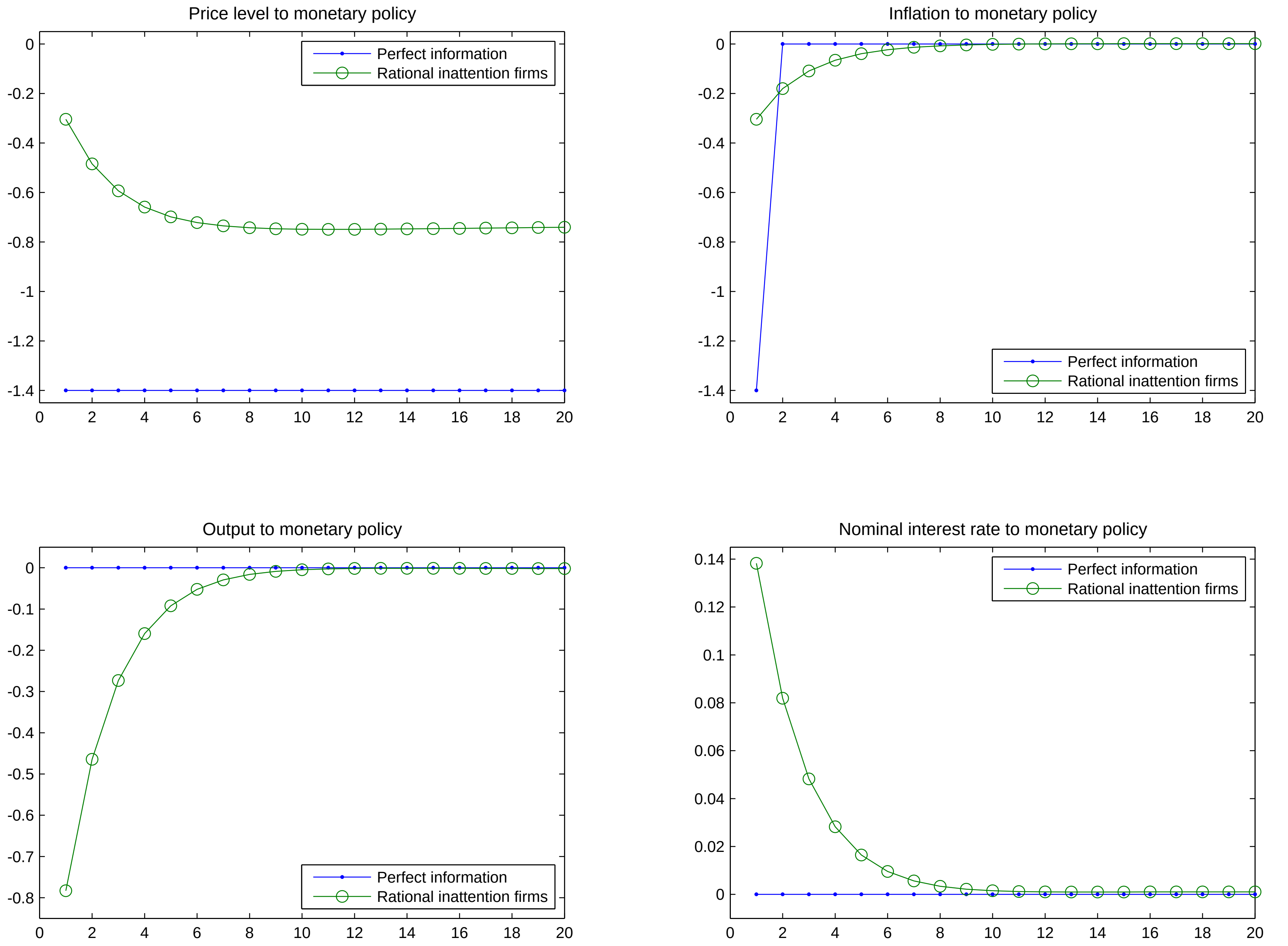


Figure 2: Impulse responses, benchmark economy

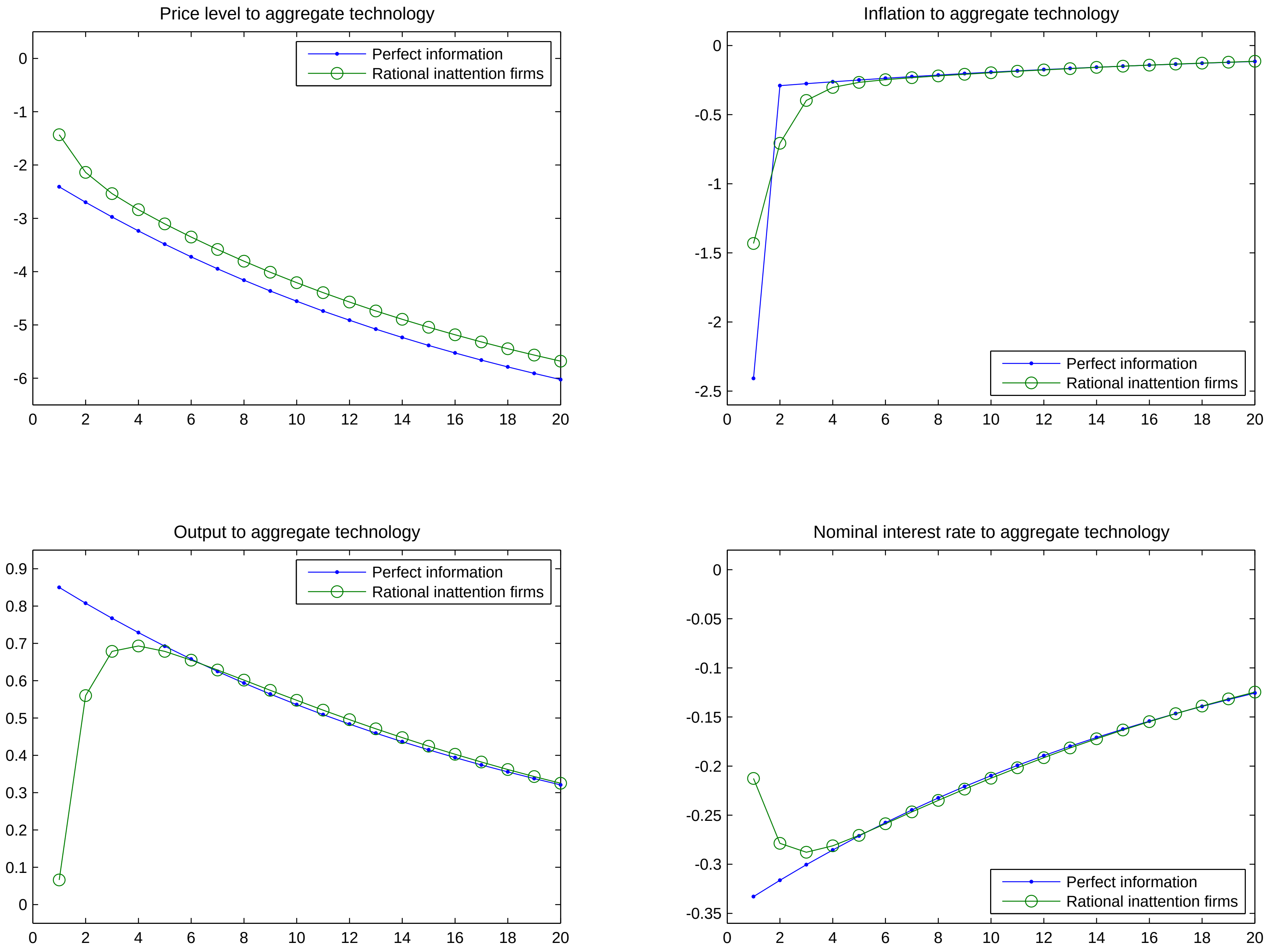


Figure 3: Impulse response of an individual price to a firm-specific productivity shock, benchmark economy

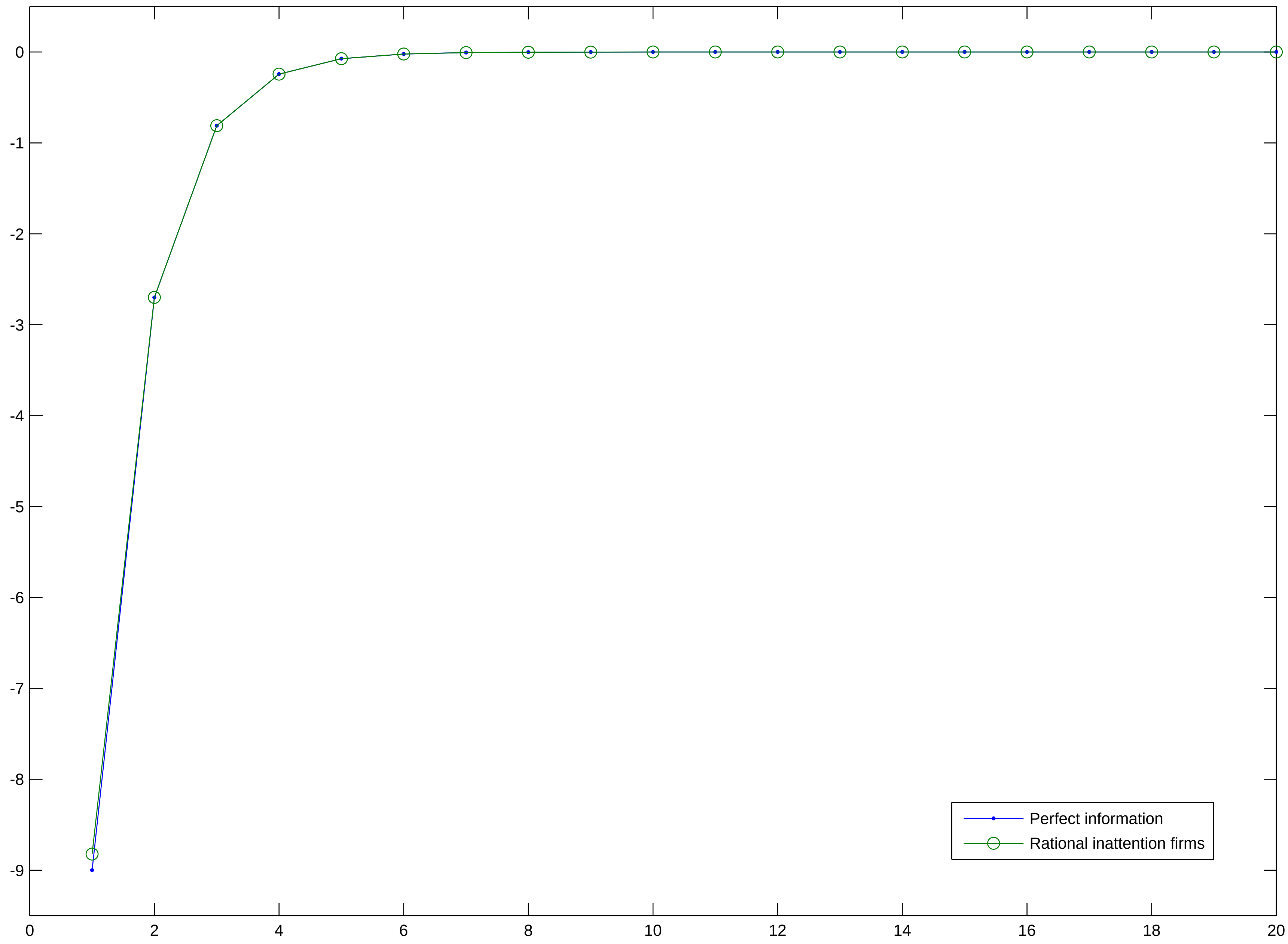


Figure 4: Impulse responses, benchmark economy and Calvo model

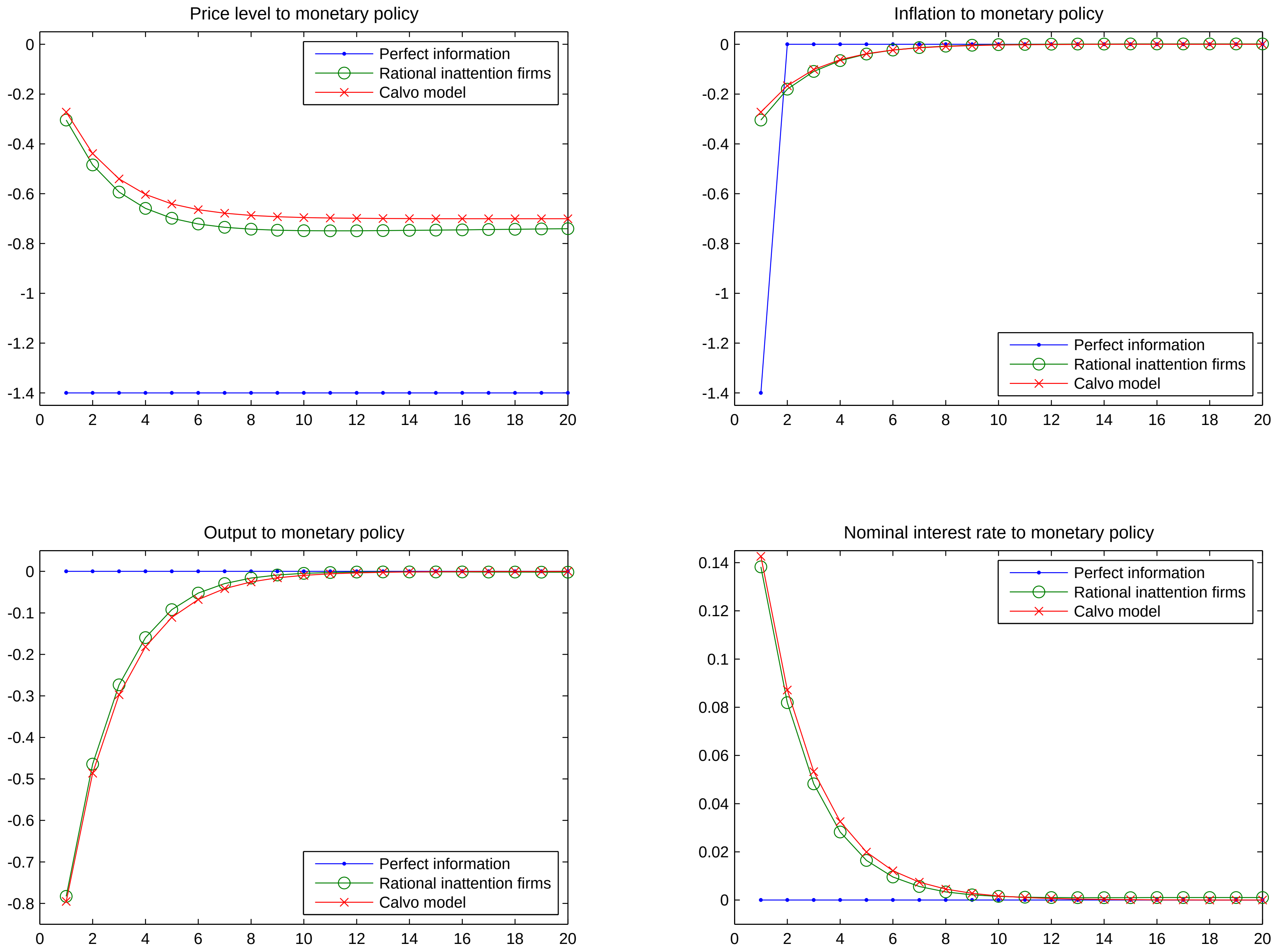


Figure 5: Impulse responses, benchmark economy and Calvo model

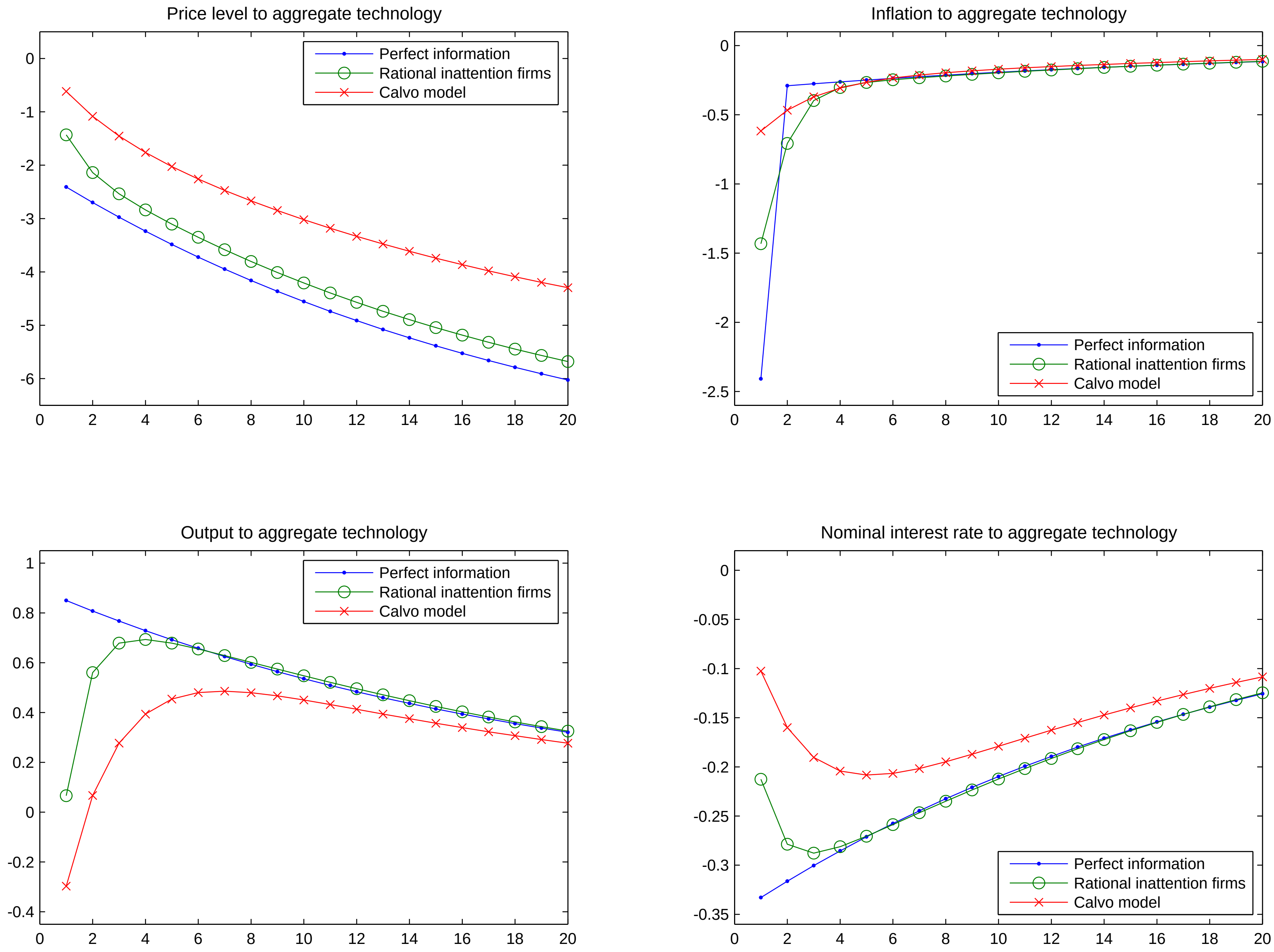


Figure 6: Standard deviation of output gap vs. parameter ϕ_π, benchmark economy and Calvo model

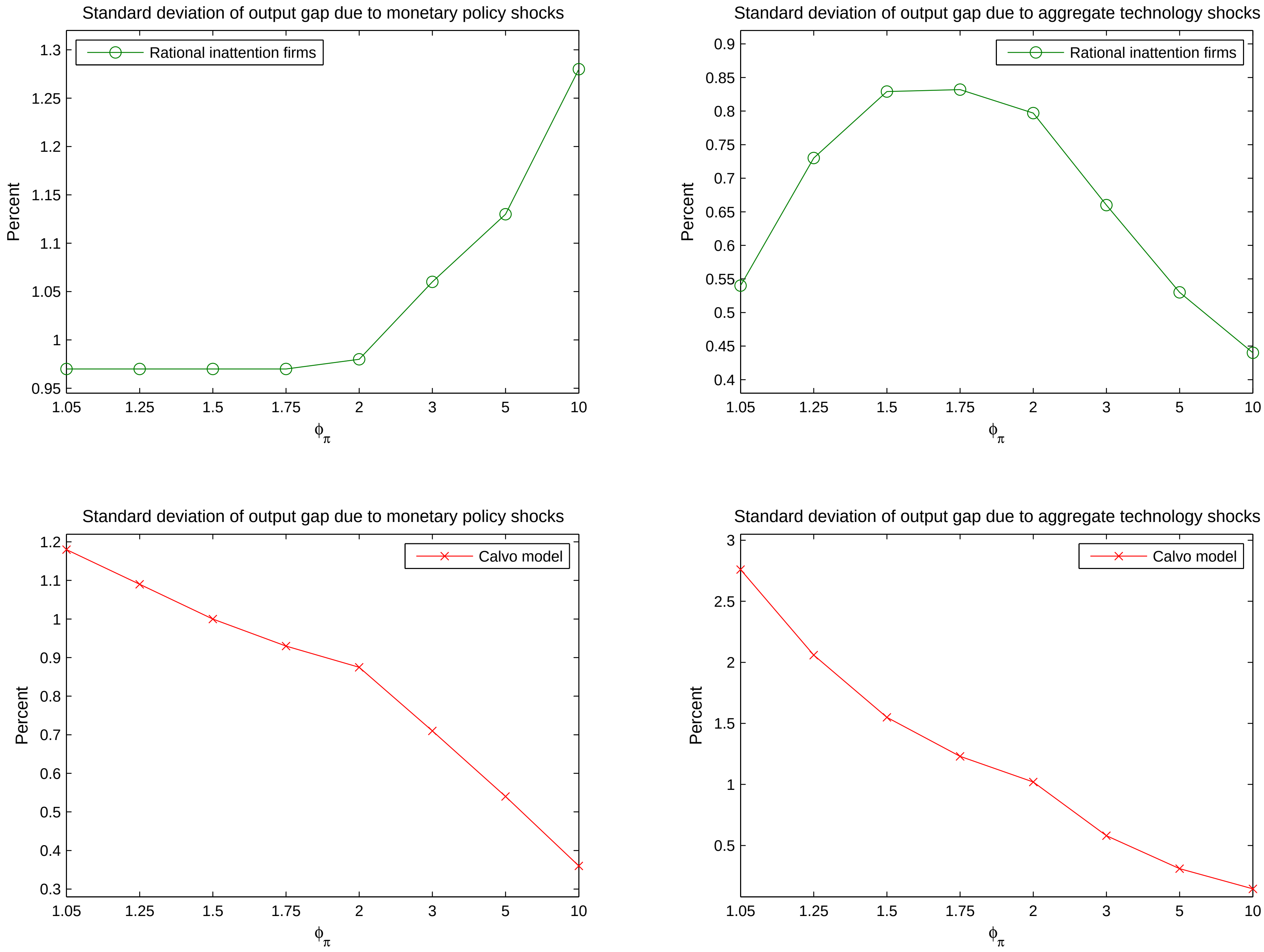
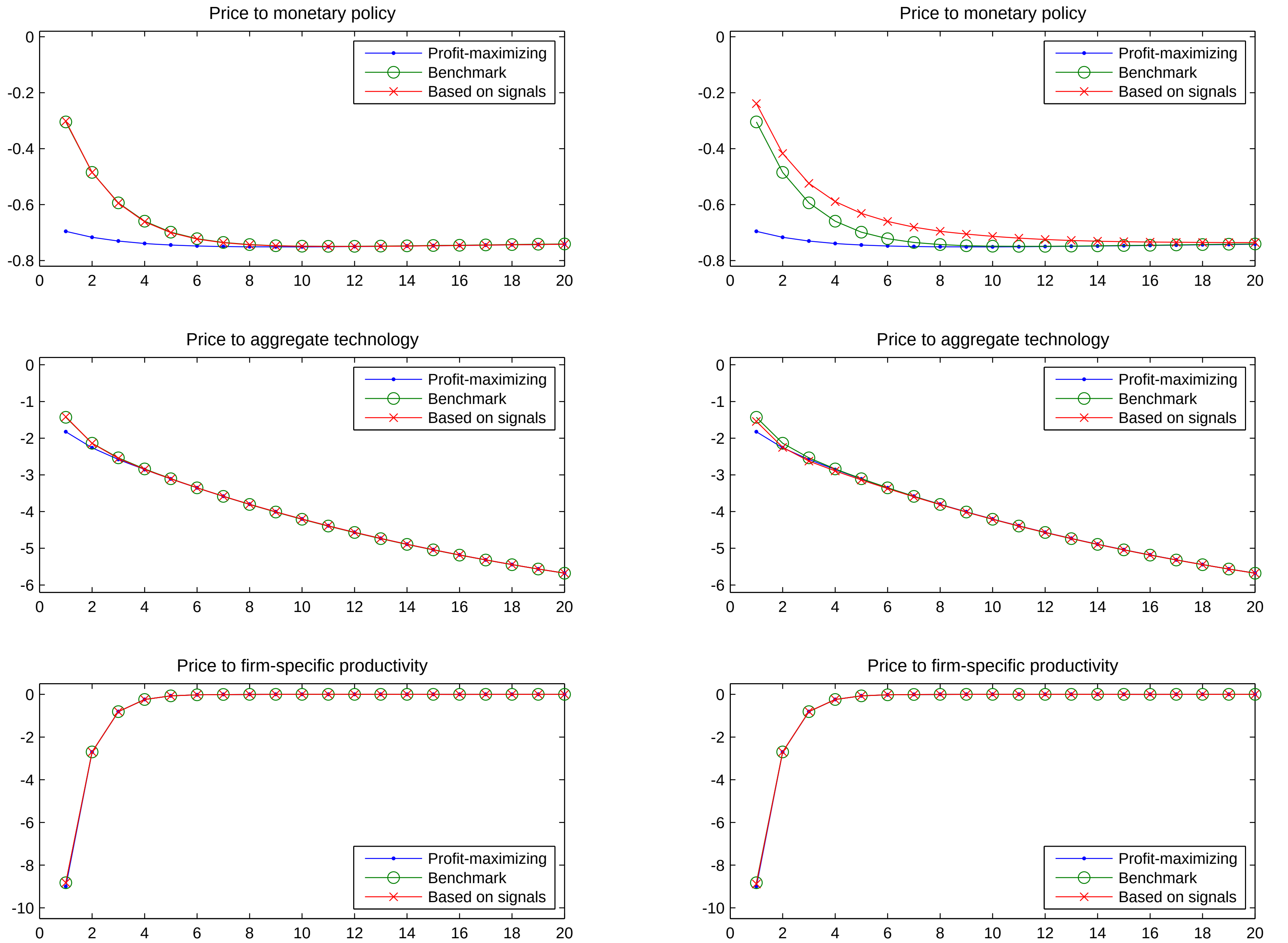


Figure 7: Impulse responses, firms' attention problem with signals



Note: Signals concerning p^{A*}, p^{R*}, and p^{I*} (left column), signals concerning the price level, TFP, last period sales, and last period wage bill (right column).

Figure 8: Impulse responses, households' problem

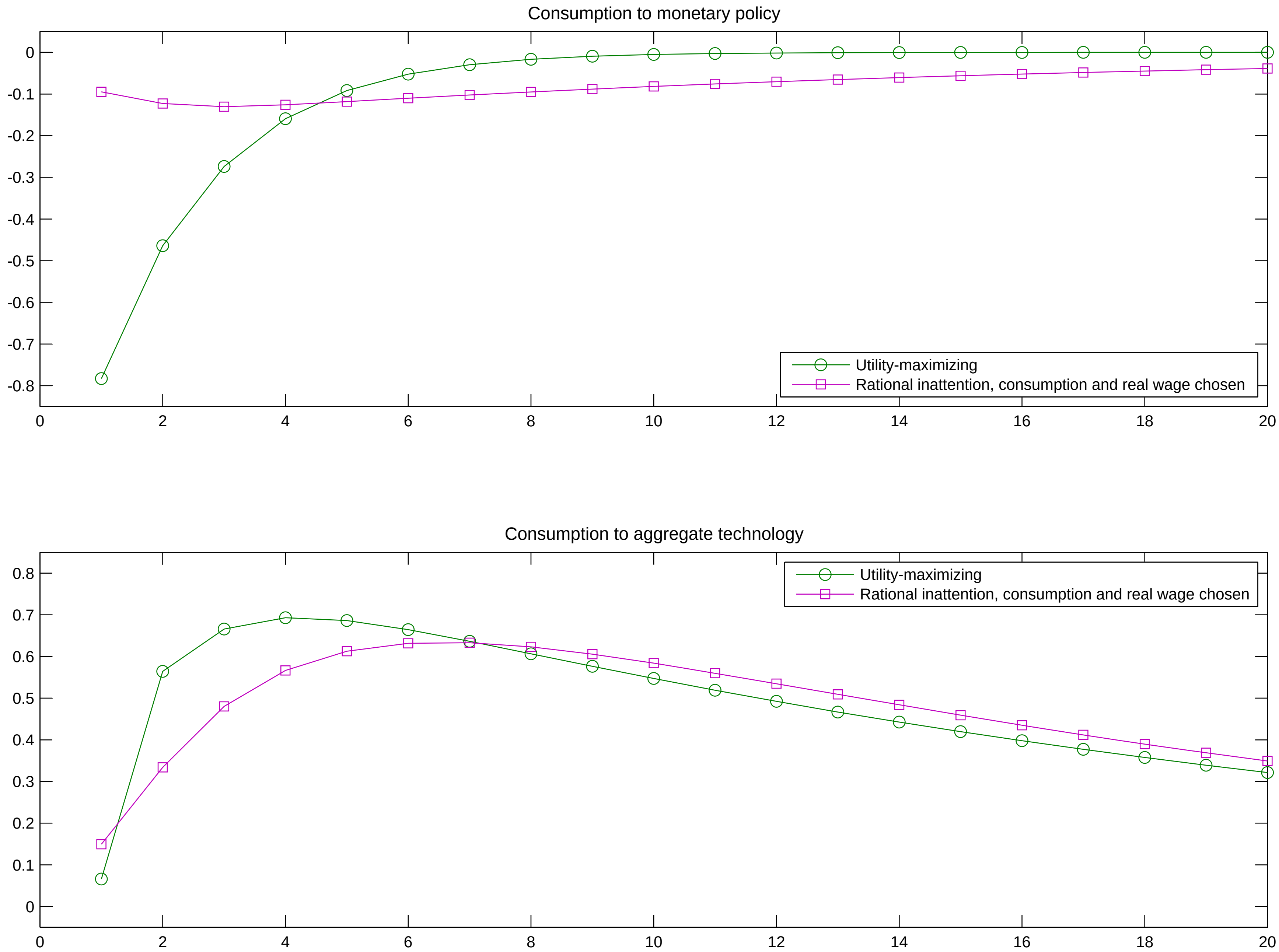


Figure 9: Impulse responses, benchmark economy

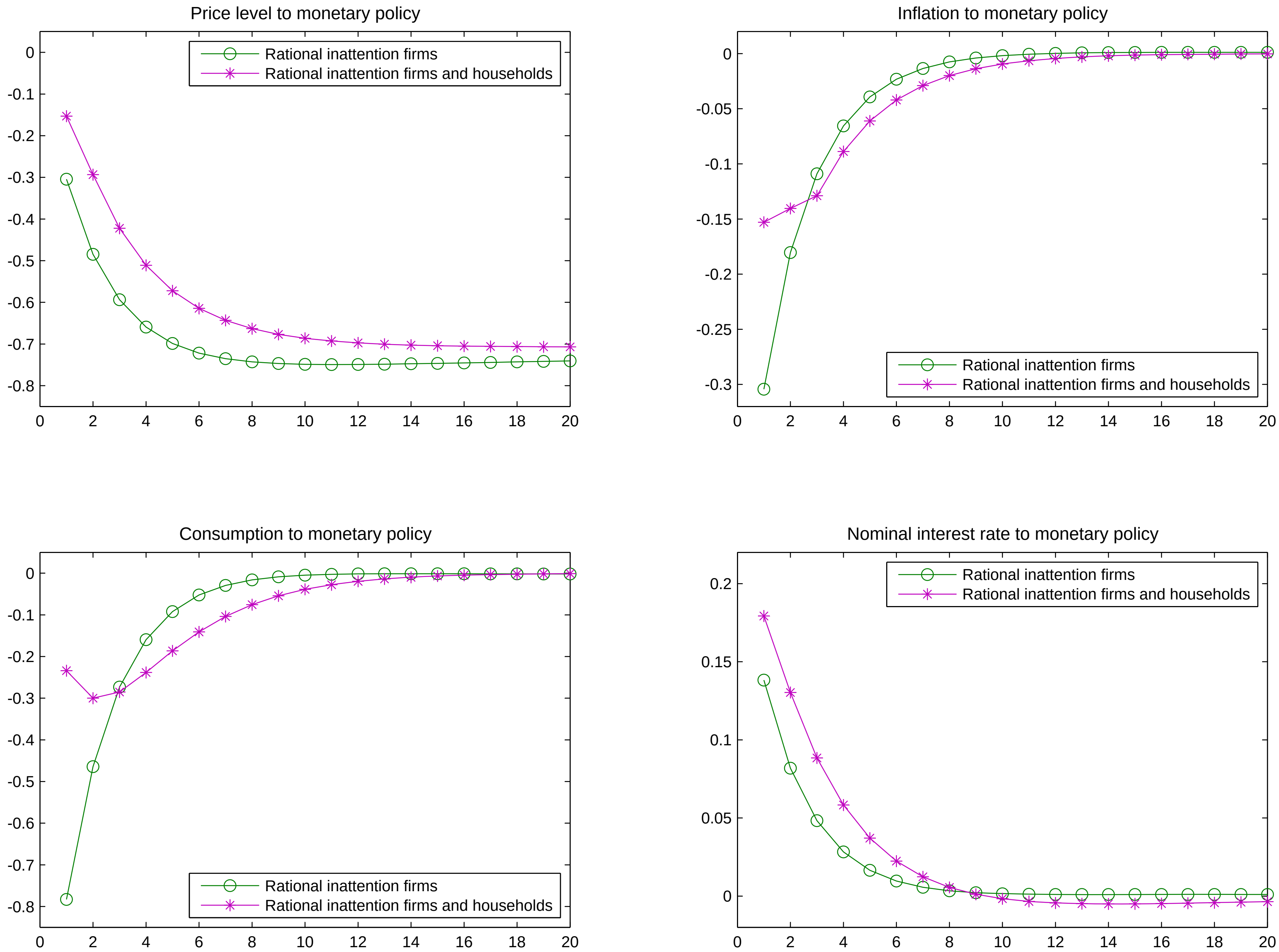


Figure 10: Impulse responses, benchmark economy