Business Cycle Dynamics under Rational Inattention*

Bartosz Maćkowiak
European Central Bank and CEPR

Mirko Wiederholt
Goethe University Frankfurt

First draft: June 2007. This draft: December 2012

Abstract

We construct and solve a dynamic stochastic general equilibrium model with rational inattention as proposed by Sims (1998). Decision-makers in firms and households have limited attention and decide how to allocate attention. The model fits macroeconomic data about as well as standard DSGE models do. At the same time, the outcomes of experiments are very different in this model than in the standard DSGE models.

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

*Maćkowiak: European Central Bank, Kaiserstrasse 29, 60311 Frankfurt am Main, Germany (e-mail: bartosz.mackowiak@ecb.int); Wiederholt: Goethe University Frankfurt, Grüneburgplatz 1, 60323 Frankfurt am Main (e-mail: wiederholt@wiwi.uni-frankfurt.de). We thank for helpful comments Bruno Biais, four anonymous referees, Paco Buera, Larry Christiano, Jim Costain, Marty Eichenbaum, Christian Hellwig, Marek Jarociński, Giorgio Primiceri, Chris Sims, Bruno Strulovici, Andrea Tambalotti as well as seminar and conference participants at Amsterdam, Bank of Canada, Bilkent, Bonn, Carlos III, Chicago Fed, Columbia, Conference in Honor of Chris Sims 2012, 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, Pompeu Fabra, Princeton, Queen Mary, Richmond Fed, Riksbank, SED 2008, Stony Brook, Toronto, Toulouse, UCSD, University of Chicago, University of Hong Kong, University of Montreal, Wharton, and Yale. We thank Jeanne Commault for excellent research assistance. 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

When Sims (1998) proposed the idea of rational inattention, his motivation was the study of business cycles. Sims considered a conventional dynamic stochastic general equilibrium model with various forms of slow adjustment of real and nominal variables. He concluded that *multiple* sources of slow adjustment were necessary for the model to match the inertia found in macroeconomic data.[1] He conjectured that the inertia in the data could instead be understood as the result of a *single* new source of slow adjustment: the assumption that people have limited attention and allocate attention optimally. He called this assumption "rational inattention."

The literature on rational inattention has grown since Sims wrote, but a DSGE model with rational inattention on the side of firms and households has not been developed yet. This paper constructs and solves a DSGE model with rational inattention on the side of firms and households. Decision-makers in firms and households have limited attention and allocate attention optimally. Following Sims (2003), limited attention is modeled as a constraint on information flow. Rational inattention is the only source of slow adjustment.

We find that the rational inattention DSGE model with a single source of slow adjustment fits macroeconomic data about as well as conventional DSGE models with multiple sources of slow adjustment. For this result, it is essential that the model include rational inattention on the side of firms *and* households. At the same time, we find that the outcomes of experiments are very different in this model than in the conventional DSGE models.

Constructing a general equilibrium model with rational inattention is a non-trivial task, because the modeler has to specify how agents who are imperfectly aware of economic conditions interact in markets. We suppose that in each market one side of the market chooses the price and the other side of the market chooses the quantity. One side of the market sets the price *having optimally allocated attention* and the other side of the market chooses the quantity *having optimally allocated attention*. In the model there are many firms, many households and a government. Firms produce differentiated goods with a variety of types of labor. Decision-makers in firms make price setting and labor mix decisions. Households consume the variety of goods, supply the differentiated types of labor and hold government bonds. Households make consumption and wage setting decisions.

[1] Later Christiano et al. (2005), Smets and Wouters (2007), Altig et al. (2011) and many others confirmed Sims's conclusion in more formal analysis.

The central bank sets the nominal interest rate according to a Taylor rule. In goods markets, firms set prices and households decide how much to buy. In labor markets, households set wage rates and firms decide how much to hire. In the bond market, the government sets the nominal interest rate and households decide how many government bonds to hold. The economy is affected by aggregate technology shocks, monetary policy shocks and firm-specific productivity shocks.

Solving for the rational expectations equilibrium of the model is also a non-trivial task, because a complicated fixed point problem arises. The optimal attention allocation by any single firm depends on the attention allocation of all other firms and all households. The optimal attention allocation by any single household depends on the attention allocation of all other households and all firms. We figure out a tractable and reliable way to solve for the rational expectations equilibrium.

To evaluate the model, we calibrate it and compare its predictions with data and with the predictions of the two most popular medium-sized DSGE models (Smets and Wouters, 2007, and Altig et al., 2011). By a comparison with data we mean: a comparison of unconditional moments in the model and in the data, and a comparison of impulse responses from the model and from the vector autoregression in Altig et al. (2011).[2]

Two salient features of the U.S. data on output and inflation are: output growth is autocorrelated and inflation is strongly autocorrelated. The model matches these two features. Rational inattention on the side of firms *and* households is essential for the model to match the two features. A version of the model with perfect information matches neither feature. Rational inattention on the side of firms yields only strongly autocorrelated inflation. The combination of rational inattention on the side of firms and households yields strongly autocorrelated inflation and autocorrelated output growth.

The rational inattention DSGE model matches the impulse responses of output to shocks about as well as the two most popular medium-sized DSGE models. The rational inattention DSGE model does slightly worse than the standard models matching the impulse response of inflation to a monetary policy shock; and does better than the standard models matching the impulse response of inflation to an aggregate technology shock.

[2] The DSGE model in Smets and Wouters (2007) and the DSGE model in Altig et al. (2011) include multiple sources of slow adjustment (Calvo price setting, habit formation in consumption, Calvo wage setting and so on) and fit the data well. The Smets-Wouters model fits the data about as well, in terms of marginal likelihood, as a VAR. The Altig et al. model was developed to fit impulse responses from a VAR. See also Christiano et al. (2005).

Let us give more details. All three models produce hump-shaped impulse responses of output to a monetary policy shock and an aggregate technology shock. All three models also yield a persistent impulse response of inflation to a monetary policy shock. However, the impulse response of inflation to a monetary policy shock is monotonic in the rational inattention DSGE model, whereas this impulse response is hump-shaped in the two standard medium-sized DSGE models and in the data. On the other hand, in the data inflation responds strongly on impact to an aggregate technology shock – an order of magnitude more strongly than to a monetary policy shock. The rational inattention DSGE model matches this fact, whereas the standard DSGE models fail to match this fact.

To develop intuition for the business cycle dynamics under rational inattention, we examine the behavior of households and firms in the DSGE model.

Consider the consumption-saving behavior of households. Utility-maximizing consumption in the model equals minus the sum of current and future real interest rates. We find that households decide to pay little attention to the real interest rate. For this reason, consumption differs significantly from utility-maximizing consumption: consumption responds *slowly* to a change in the real interest rate. This result is important because in a large class of models monetary policy affects the real economy through the following channel. The central bank moves the nominal interest rate; due to some form of price stickiness the real interest rate changes; and consumption responds to the change in the real interest rate. The model predicts that the last part of this channel will be slow, that is, consumption will respond slowly to a change in the real interest rate.

In the data, consumption also responds slowly to a change in the real interest rate. The literature on VARs finds that consumption shows a hump-shaped response to a monetary policy shock.[3] The literature on conventional DSGE models finds that the fit of those models to macroeconomic data is maximized when the degree of habit formation in consumption is large.[4] With a large degree of habit formation, consumption responds slowly to a change in the real interest rate. Our model suggests that the observed slow reaction of consumption to the real interest rate is the outcome of a decision problem by households with standard preferences but with limited attention.

The result that households choose to pay little attention to the real interest rate holds for low

[3] See, for example, Leeper et al. (1996).

[4] See, for example, Smets and Wouters (2007), Justiniano and Primiceri (2008) and Altig et al. (2011).

and high values of the coefficient of relative risk aversion. When risk aversion is low, a deviation of consumption from utility-maximizing consumption costs little in utility terms. As a result, it is unimportant to be aware of movements in the real interest rate. When risk aversion is high, utility-maximizing consumption varies little with a change in the real interest rate. Therefore, it is again unimportant to be aware of movements in the real interest rate.

Next, consider the price setting behavior of firms. We find that decision-makers in firms decide to allocate little attention to monetary policy, some attention to aggregate technology and a lot of attention to firm-specific productivity. This allocation of attention implies that prices respond slowly to monetary policy shocks, fairly quickly to aggregate technology shocks and very quickly to firm-specific productivity shocks. The endogenous attention allocation allows the model to: (i) produce a much stronger response of inflation to an aggregate technology shock than to a monetary policy shock, consistent with the aggregate data, and (ii) match the empirical finding by Boivin et al. (2009) and Maćkowiak et al. (2009) that prices respond very quickly to disaggregate shocks.

Furthermore, the endogenous attention allocation implies that the model yields strong and persistent real effects of monetary policy shocks while *profit losses are small*. In any model with a price setting friction, firms experience profit losses due to deviations of the price from the profit-maximizing price. In our calibrated model, the expected per-period loss in profit due to deviations of the price from the profit-maximizing price is *thirty-five times smaller* than the analogous number in a version of the Calvo model producing the same real effects of monetary policy shocks. The reason why profit losses are so much smaller in the rational inattention model is that in this model prices respond slowly only to unimportant shocks but quickly to important shocks.

We use the model to conduct experiments. We find that the outcomes of experiments are very different in this model than in the conventional DSGE models currently used to analyze monetary policy, even though the models yield similar impulse responses. Moreover, there is a systematic reason for why the outcomes are so different: the allocation of attention varies with the environment.

Consider perhaps the most common experiment in the literature on DSGE models used for analysis of monetary policy, a change in the coefficient on inflation in the Taylor rule. In simple New Keynesian models and in models with exogenous dispersed information, there is a *monotonic* negative relationship between the coefficient on inflation in the Taylor rule and the variance of the output gap due to aggregate technology shocks. By contrast, in our model this relationship is *non-*

monotonic. The reason is that in our 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. This effect per se makes the output gap *more* volatile. We find that, for reasonable parameter values, this additional effect dominates and fighting inflation more aggressively *increases* the variance of the output gap due to aggregate technology shocks. This is important: the rational inattention DSGE model gives a very different answer than the conventional DSGE models to the basic question what happens to the real economy when monetary policy fights inflation more aggressively.

Another conventional wisdom derived from DSGE models currently used for monetary policy analysis is that more strategic complementarity in price setting implies stronger real effects of monetary policy shocks. A common way to raise strategic complementarity in price setting is to make a firm's marginal cost curve more upward sloping in own output.[5] When we increase strategic complementarity in price setting by making a firm's marginal cost curve more upward sloping in own output, we find that, for reasonable parameter values, real effects of monetary policy shocks become *smaller*. The reason is that in our 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. Decision-makers in firms therefore decide to pay more attention to the price setting decision, implying that prices respond faster to shocks. This additional effect dominates for reasonable parameter values and thus real effects of monetary policy shocks become *smaller*. The conventional wisdom that strategic complementarity in price setting necessarily strengthens real effects of monetary policy does not hold in the rational inattention DSGE model.

Let us also describe what happens when we raise the standard deviation of any aggregate shock. Decision-makers in firms and households decide to pay more attention to the aggregate economy. This result explains the evidence in Coibion and Gorodnichenko (2012) who study survey data on expectations finding that the degree of attention to the aggregate economy rose markedly during the turbulent 1970s and fell significantly during the subsequent calm period.

This paper belongs to the literature on rational inattention following Sims (2003).[6] The main

[5] See, for example, Altig et al. (2011).

[6] See, for example, Sims (2006), Kacperczyk et al. (2012), Luo (2008), Maćkowiak and Wiederholt (2009), Matejka (2011), Matejka and Sims (2010), Mondria (2010), Paciello (2012), Paciello and Wiederholt (2012), Van Nieuwerburgh and Veldkamp (2009, 2010) and Woodford (2009). Sims (2010), Veldkamp (2011) and Wiederholt (2010) review the

difference to all the existing literature on rational inattention is that we solve a DSGE model with rational inattention on the side of firms and households, and that we compare the model quantitatively with data and conventional DSGE models. The most closely related paper is Maćkowiak and Wiederholt (2009). In that paper the demand side of the economy is an exogenous process for nominal spending, whereas here the demand side of the economy is determined by households' optimization and a monetary policy rule. This allows us to conduct experiments that central banks are interested in (e.g., what happens when the central bank fights inflation more aggressively). Moreover, households optimize under rational inattention. This is critical for the ability of the model to match the data.[7]

This paper studies consumption by households with limited attention when the real interest rate fluctuates. Sims (2003, 2006), Luo (2008) and Tutino (2012) also study consumption-saving decisions under rational inattention but assume that the real interest rate is constant.[8] Therefore, the point that households have little incentive to attend to the real interest rate (for low and high values of the coefficient of relative risk aversion) is not in those papers. This point is important because in a large class of models monetary policy affects the real economy by moving the real interest rate. If this is indeed the channel through which monetary policy affects the real economy, then the attention that households devote to the real interest rate is crucial.

The paper is also related to the literature on business cycle models with imperfect information (e.g., Lucas (1972), Woodford (2002), Mankiw and Reis (2002), Angeletos and La'O (2009a, 2009b) and Lorenzoni (2009)). The main difference to this literature is that in our model decision-makers choose the information structure, i.e., the information structure is derived from an objective and constraints. This has two implications. First, the model gives an explanation for the equilibrium information structure. Second, the model predicts how the equilibrium information structure varies with policy. The fact that the equilibrium information structure varies with policy has important implications for the outcomes of experiments.

literature on rational inattention.

[7] Paciello (2012) solves a stochastic general equilibrium model with rational inattention on the side of firms. The main differences are that in his model households have perfect information 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.

[8] The real interest rate is constant also in Reis (2006a).

Section 2 describes all features of the model apart from the attention problems of firms and households. Section 3 describes the attention problems. Section 4 solves the model and evaluates how well the model matches the data. Section 5 uses the model to conduct experiments. Section 6 considers extensions in which we vary assumptions concerning information flows. Section 7 concludes.

2 Model setup – physical environment

In this section we describe preferences and technology, market structure and asset structure, and monetary and fiscal policy. These features of the economy are almost identical to a simple New Keynesian model, apart from the fact that we discard *all* sources of slow adjustment that usually are in New Keynesian models (Calvo pricing, habit formation in consumption, Calvo wage setting).[9] In the next section we describe how decision-makers in firms and households make decisions under rational inattention. Rational inattention will be the *only* source of slow adjustment.

Time is discrete. There are three types of markets: goods markets, labor markets and a government bond market. In each market, one side of the market sets the price and the other side of the market chooses the quantity. In goods markets, firms set prices and households decide how much to buy. In labor markets, households set wage rates and firms decide how much to hire. In the bond market, the government sets the nominal interest rate and households decide how many government bonds to hold. This setup is convenient for formulating a DSGE model with rational inattention and happens to be the setup of a standard New Keynesian model.

2.1 Households

There are J households. Households consume a variety of goods, supply labor and hold government bonds. Since households supply differentiated types of labor, they have market power in the labor market.

Households have an infinite horizon. Each household seeks to maximize the expected discounted

[9] For reasons explained below, we make two additional changes to the standard New Keynesian model. We assume that there is a large finite number of firms (instead of a continuum of firms) and that asset markets are incomplete.

sum of period utility. The discount factor is $\beta \in (0,1)$. The period utility function is

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

with

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

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

Households can trade a single asset: nominal government bonds.[10] 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 R_{t-1} is the gross nominal interest rate on bond holdings between period $t-1$ and period t, B_{jt-1} are bond holdings by household j between period $t-1$ and period t, τ_w is a wage subsidy, W_{jt} is the nominal wage rate for labor supplied by household j in period t, (D_t/J) is a pro-rata share of nominal aggregate profits, (T_t/J) is a pro-rata share of nominal lump-sum taxes, and P_{it} is the price of good i in period t. Each household has the same initial bond holdings. To rule out Ponzi schemes, we assume that bond holdings have to be positive in every period, $B_{jt} > 0$.[11]

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: prices of consumption goods, the nominal interest rate, the aggregate wage index defined below, and all aggregate quantities.

[10] We assume that asset markets are incomplete, because the assumption of complete markets seems even stronger than usual in an environment where agents are imperfectly aware of economic conditions.

[11] One has to make some assumption to rule out Ponzi schemes. We choose this particular assumption because it allows 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.

2.2 Firms

There are I firms. Firms supply differentiated goods. Firm i supplies good i. The production function of firm i is

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

with

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

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

Nominal profit of firm i in period t equals revenue minus cost

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

where τ_p is a production subsidy paid by the government.

In every period, each firm sets a price, P_{it}, and chooses a labor mix, $\left(\hat{L}_{i1t}, \ldots, \hat{L}_{i(J-1)t}\right)$. Here $\hat{L}_{ijt} = (L_{ijt}/L_{it})$ 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 and produces the quantity demanded with the chosen labor mix.

Each firm takes as given: the consumer price index defined below, the nominal interest rate, wage rates, all aggregate quantities, and total factor productivity.[12]

2.3 Government

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

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

[12] Dixit and Stiglitz (1977) also assume that there is a finite number of firms and firms take the consumer price index as given. This seems a good description of the U.S. economy.

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

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

Y_t^P is potential output and ε_t^R is a monetary policy shock. The price index P_t and potential output are defined later. R and Π denote the values of the nominal interest rate and inflation, respectively, in the non-stochastic steady state. The policy parameters satisfy $\rho_R \in [0,1)$, $\phi_\pi > 1$ and $\phi_y \geq 0$.

The government budget constraint in period t reads

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

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

Following common practice in the New Keynesian literature, we assume that the government sets the production subsidy τ_p so as to correct the distortion arising from firms' market power in the goods market, and the government sets the wage subsidy τ_w so as to correct the distortion arising from households' market power in the labor market. Formally,

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

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

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

where $\tilde{\eta}$ denotes the wage elasticity of labor demand.[13]

Following again common practice in the New Keynesian literature, we assume that monetary policy is active and fiscal policy is passive in the sense of Leeper (1991).

2.4 Shocks

There are three types of shocks: monetary policy shocks, aggregate technology shocks and firm-specific productivity shocks. The stochastic processes $\{\varepsilon_t^R\}$, $\{a_t\}$, and $\{a_{1t}\}$, $\{a_{2t}\}$,..., $\{a_{It}\}$ are

[13] The price elasticity of demand $\tilde{\theta}$ equals the preference parameter θ when households have perfect information. However, when households have imperfect information the price elasticity of demand $\tilde{\theta}$ may differ from the preference parameter θ. Therefore, the value of the production subsidy given by equation (10) may vary across information structures. An analogous comment applies to the wage subsidy.

independent. The variable ε_t^R follows a Gaussian white noise process, a_t follows a stationary Gaussian first-order autoregressive process with mean zero, and each a_{it} also follows a stationary Gaussian first-order autoregressive process with mean zero. 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 small by setting the number of firms I equal to a large number. We work with a finite number of firms because a household subject to rational inattention cannot track a continuum of prices.

2.5 Notation

Before proceeding to the next section, it is useful to introduce notation. Throughout the paper, C_t denotes aggregate composite consumption and L_t denotes aggregate composite labor input. Formally,

$$C_t = \sum_{j=1}^{J} C_{jt}, \qquad L_t = \sum_{i=1}^{I} L_{it}. \tag{12}$$

Furthermore, $\hat{P}_{it}$ denotes the relative price of good i and $\hat{W}_{jt}$ denotes the relative wage rate for type j labor. Formally,

$$\hat{P}_{it} = \frac{P_{it}}{P_t}, \qquad \hat{W}_{jt} = \frac{W_{jt}}{W_t}.$$

We specify the definitions of P_t and W_t later. Finally, $\tilde{W}_{jt}$ denotes the real wage rate for type j labor and $\tilde{W}_t$ denotes the real wage index. Formally,

$$\tilde{W}_{jt} = \frac{W_{jt}}{P_t}, \qquad \tilde{W}_t = \frac{W_t}{P_t}.$$

3 Model setup – rational inattention

We now describe how decision-makers in firms and households allocate attention. A decision-maker who allocates attention optimally compares cost and benefit of paying attention. The benefit of paying attention to the current state of the economy (e.g., interest rates) is that decisions get closer to the optimal decisions under perfect information. The cost of paying attention can be thought of as time. Paying attention uses up some of the agent's valuable time.

To evaluate the benefit of paying attention for a firm in our model, we derive a simple expression for the loss in profit that a firm incurs if the firm takes actions that deviate from the optimal actions under perfect information. We then state the attention problem of the decision-maker in a firm.

To evaluate the benefit of paying attention for a household, we derive a simple expression for the loss in utility that a household incurs if the household takes actions that deviate from the optimal actions under perfect information. We then state the attention problem of the household.

Finally, we explain how decisions by individual firms and households are aggregated.

3.1 Loss in profit in the case of suboptimal actions

We derive an expression for the loss in profit that a firm incurs if the firm chooses a price and a labor mix that deviate from the optimal decisions under perfect information. This expression is derived from the technology presented in Section 2, the requirement that a firm meets demand and the demand function for good i.

We guess the following demand function

$$C_{it} = \vartheta \left(\frac{P_{it}}{P_t} \right)^{-\tilde{\theta}} C_t. \tag{13}$$

We verify below that the equilibrium demand function has this form. Here C_{it} is demand for good i in period t, C_t is aggregate composite consumption, P_{it} is the price of good i, and $P_t = d(P_{1t}, \ldots, P_{It})$ is a price index. The function d is homogenous of degree one, continuously differentiable and symmetric. The coefficients $\tilde{\theta}$ and ϑ satisfy $\tilde{\theta} > 1$ and $\vartheta = I^{-\frac{\theta - \tilde{\theta}}{\theta - 1}}$.

Substituting the production function (4)-(5), the requirement that output equals demand and the demand function (13) into the expression for profit (6) yields

$$(1+\tau_p) P_{it} \vartheta \left(\frac{P_{it}}{P_t} \right)^{-\tilde{\theta}} C_t - \left[\frac{\vartheta \left(\frac{P_{it}}{P_t} \right)^{-\tilde{\theta}} C_t}{e^{a_t} e^{a_{it}}} \right]^{\frac{1}{\alpha}} \left[\sum_{j=1}^{J-1} W_{jt} \hat{L}_{ijt} + W_{Jt} \left(1 - \sum_{j=1}^{J-1} \hat{L}_{ijt}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \right]. \tag{14}$$

Profit equals revenue minus cost. Cost is expressed as the product of composite labor input and the wage bill per unit of composite labor input. Recall that $\hat{L}_{ijt} = (L_{ijt}/L_{it})$ is firm i's relative input of type j labor.

The economy described in Section 2 is an incomplete markets economy with multiple owners of a firm. It is therefore in principle unclear how firms value profit in different states of the world.

For this reason, we assume a general stochastic discount factor. In particular, we assume that decision-makers in firms in period -1 value nominal profit in period t using

$$Q_{-1,t} = \beta^t \Lambda\left(C_{1t}, \ldots, C_{Jt}\right) \frac{1}{P_t}, \tag{15}$$

where Λ is some twice continuously differentiable function and P_t is the price index appearing in the demand function (13). Then, $Q_{-1,t}$ times nominal profit in period t equals

$$\begin{aligned} & \beta^t \Lambda\left(C_{1t}, \ldots, C_{Jt}\right)\left(1+\tau_p\right) \vartheta \hat{P}_{it}^{1-\tilde{\theta}} \left(\sum\nolimits_{j=1}^{J} C_{jt}\right) \\ -\beta^t \Lambda\left(C_{1t}, \ldots, C_{Jt}\right) & \left[\frac{\vartheta \hat{P}_{it}^{-\tilde{\theta}} \left(\sum\nolimits_{j=1}^{J} C_{jt}\right)}{e^{a_t} e^{a_{it}}}\right]^{\frac{1}{\alpha}} \left[\sum_{j=1}^{J-1} \tilde{W}_{jt} \hat{L}_{ijt} + \tilde{W}_{Jt} \left(1 - \sum_{j=1}^{J-1} \hat{L}_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\right], \end{aligned} \tag{16}$$

where $\hat{P}_{it} = \left(P_{it}/P_t\right)$ is the relative price of good i and $\tilde{W}_{jt} = \left(W_{jt}/P_t\right)$ is the real wage rate for type j labor.

Computing a log-quadratic approximation to the real profit function (16) around the non-stochastic steady state yields a simple expression for the loss in profit that a firm incurs if the firm chooses a price and a labor mix that deviate from the optimal decisions under perfect information. 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.[14] For example, $\hat{p}_{it} = \ln\left(\hat{P}_{it}/\hat{P}_i\right)$ is the log deviation of the relative price of good i from its value in the non-stochastic steady state. Let x_t denote the vector of variables appearing in the real profit function (16) that the firm can affect, i.e., the price of good i and the labor mix

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

Let $E_{i,-1}$ denote the expectation operator conditioned on the information of the decision-maker in firm i in period -1. After the log-quadratic approximation to the real profit function (16), the expected discounted sum of losses in profit in the case of suboptimal actions equals

$$\sum_{t=0}^{\infty} \beta^t E_{i,-1}\left[\frac{1}{2}\left(x_t - x_t^*\right)' H \left(x_t - x_t^*\right)\right], \tag{17}$$

[14] It is straightforward to solve for the non-stochastic steady state of the economy presented in Section 2. See Appendix A in Maćkowiak and Wiederholt (2010).

where the matrix H is given by

$$H=-\Lambda\left(C_1,\ldots,C_J\right)\tilde{W}L_i\begin{bmatrix}\frac{\tilde{\theta}}{\alpha}\left(1+\frac{1-\alpha}{\alpha}\tilde{\theta}\right) & 0 & \cdots & \cdots & 0\\ 0 & \frac{2}{\eta J} & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J}\\ \vdots & \frac{1}{\eta J} & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & \frac{1}{\eta J}\\ 0 & \frac{1}{\eta J} & \cdots & \frac{1}{\eta J} & \frac{2}{\eta J}\end{bmatrix}, \tag{18}$$

and the vector of optimal decisions under perfect information x_t^* is given by

$$\hat{p}_{it}^*=\frac{\frac{1-\alpha}{\alpha}}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\left(\frac{1}{J}\sum_{j=1}^{J}c_{jt}\right)+\frac{1}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\left(\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}\right)-\frac{\frac{1}{\alpha}}{1+\frac{1-\alpha}{\alpha}\tilde{\theta}}\left(a_t+a_{it}\right), \tag{19}$$

and

$$\hat{l}_{ijt}^*=-\eta\left(\tilde{w}_{jt}-\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt}\right). \tag{20}$$

In summary, after the log-quadratic approximation to the real profit function, the optimal decisions under perfect information are given by the usual log-linear equations (19)-(20). Furthermore, for any process $\{x_t\}_{t=0}^{\infty}$, the expected discounted sum of losses in profit due to deviations of the price and the labor mix from the optimal decisions under perfect information is given by expression (17).[15] See Proposition 1 in Maćkowiak and Wiederholt (2010).

3.2 Attention problem of the decision-maker in a firm

From the last subsection we know how much profit a firm loses if the firm takes actions that differ from the optimal actions under perfect information. We now state the attention problem of the decision-maker in the firm. We assume that the decision-maker allocates his attention so as to maximize expected profit of the firm net of the cost of paying attention. The only friction is that paying attention is costly. If attention were free, the decision-maker would take the optimal actions under perfect information.

In the rational inattention literature, there are two approaches to stating the attention problem. One approach is to let agents choose the precision of signals, subject to the constraint that the

[15] To be precise, in the derivation of expression (17) we use a technical condition stated in Proposition 1 in Maćkowiak and Wiederholt (2010). This technical condition places some restrictions on the process $\{x_t\}_{t=0}^{\infty}$. This technical condition is satisfied in the paper.

signals can only contain a limited amount of information. After agents have chosen the precision of signals, they receive the signals and take optimal actions given the signals. Another approach is to let agents choose directly a law of motion for their actions, subject to the constraint that the *actions* can only contain a limited amount of information. In Section V of Maćkowiak and Wiederholt (2009) we establish a formal equivalence of these two approaches under certain conditions. Here we follow the second approach. We let agents choose directly a law of motion for their actions, subject to a constraint on the information content of the actions.[16]

The decision-maker in firm i solves the following attention problem in period -1:

$$\max_{\kappa,B(L),C(L),\tilde{\eta},\chi}\left\{\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}\left(x_t-x_t^*\right)' H\left(x_t-x_t^*\right)\right]-\frac{\mu}{1-\beta}\kappa\right\}, \tag{21}$$

where

$$x_t-x_t^*=\begin{pmatrix} p_{it} \\ \hat{l}_{i1t} \\ \vdots \\ \hat{l}_{i(J-1)t}\end{pmatrix}-\begin{pmatrix} p_{it}^* \\ \hat{l}_{i1t}^* \\ \vdots \\ \hat{l}_{i(J-1)t}^*\end{pmatrix}, \tag{22}$$

subject to the law of motion for the optimal actions under perfect information

$$p_{it}^*=\underbrace{A_1\left(L\right)\varepsilon_t^A}_{p_{it}^{*A}}+\underbrace{A_2\left(L\right)\varepsilon_t^R}_{p_{it}^{*R}}+\underbrace{A_3\left(L\right)\varepsilon_{it}^I}_{p_{it}^{*I}}, \tag{23}$$

$$\hat{l}_{ijt}^*=-\eta\hat{w}_{jt}, \tag{24}$$

the law of motion for the actions

$$p_{it}=\underbrace{B_1\left(L\right)\varepsilon_t^A+C_1\left(L\right)\nu_{it}^A}_{p_{it}^A}+\underbrace{B_2\left(L\right)\varepsilon_t^R+C_2\left(L\right)\nu_{it}^R}_{p_{it}^R}+\underbrace{B_3\left(L\right)\varepsilon_{it}^I+C_3\left(L\right)\nu_{it}^I}_{p_{it}^I}, \tag{25}$$

$$\hat{l}_{ijt}=-\tilde{\eta}\left(\hat{w}_{jt}+\frac{Var\left(\hat{w}_{jt}\right)}{\chi}\nu_{ijt}^L\right), \tag{26}$$

and the constraint on the information content of the actions

$$\mathcal{I}\left(\left\{p_{it}^{*A},p_{it}^{*R},p_{it}^{*I},\hat{l}_{i1t}^*,\ldots,\hat{l}_{i(J-1)t}^*\right\};\left\{p_{it}^A,p_{it}^R,p_{it}^I,\hat{l}_{i1t},\ldots,\hat{l}_{i(J-1)t}\right\}\right)\leq\kappa. \tag{27}$$

[16] In Section 6 we show that we obtain the same results when we let decision-makers in firms choose the precision of signals, subject to a constraint on the information content of the signals.

To understand problem (21)-(27), begin by considering price setting of the firm. The first element of $x_t - x_t^*$ in equation (22) is the deviation of the price from the profit-maximizing price. The loss in profit due to a suboptimal price in period t is given by the quadratic form in objective (21). If the loss in profit is larger, the quadratic form is more negative. The profit-maximizing price in period t is given by equation (23). We guess that the profit-maximizing price is a linear function of current and past shocks, a guess that we verify below. $A_1(L)$ to $A_3(L)$ are infinite-order lag polynomials. For example, the lag polynomial $A_1(L)$ gives the response of the profit-maximizing price to aggregate technology shocks. The decision-maker in the firm chooses the law of motion for the price. See equation (25). $B_1(L)$ to $B_3(L)$ and $C_1(L)$ to $C_3(L)$ are infinite-order lag polynomials. For example, if the decision-maker chooses $B_1(L) = A_1(L)$ and $C_1(L) = 0$, he chooses to respond perfectly to aggregate technology shocks. The variables ν_{it}^A, ν_{it}^R, and ν_{it}^I are information-processing errors. These variables follow Gaussian white noise processes that are independent of fundamentals, independent across firms, independent of each other, and have unit variance. Choosing the law of motion for the price is like choosing the precision of a signal about aggregate technology, the precision of a signal about monetary policy and the precision of a signal about firm-specific productivity. A more precise signal leads to a response to fundamentals that is closer to the profit-maximizing response and changes the amount of noise in the pricing behavior. In Section 6, we show that the agent's price setting behavior is *identical* when the agent chooses the precision of signals about the fundamentals in period -1, receives the noisy signals in the following periods, and sets the optimal price given the signals in every period.

The decision-maker in the firm would like to set a price equal to the profit-maximizing price in every period. If attention were free, this is what the decision-maker would do. However, we assume that paying attention is costly. Formally, the decision-maker faces constraint (27). The left-hand side of constraint (27) measures the amount of information acquired and processed by the agent. Following Sims (2003), the amount of information acquired and processed by the agent is measured by the amount of information contained in the agent's actions about the optimal actions under perfect information. Information contained in one variable about another variable is quantified as uncertainty reduction, where uncertainty is measured by entropy. The operator $\mathcal{I}$ is defined in Appendix A. Constraint (27) says that if the decision-maker wants to take more informed actions (i.e., if the decision-maker chooses a law of motion for the actions that is associated with a larger left-hand side of the constraint), he has to pay more attention (i.e., he has to choose a larger κ). Objective (21) states that paying attention is costly. The variable κ is the amount of attention devoted to the price setting decision and the labor mix decision. The parameter $\mu > 0$ is the per-period marginal cost of paying attention.

The way we model the attention problem of the decision-maker in a firm is almost identical to the way we model it in Maćkowiak and Wiederholt (2009). There are three differences. First, in Maćkowiak and Wiederholt (2009) there are only two types of shocks: nominal shocks and idiosyncratic shocks. Here there are aggregate technology shocks, monetary policy shocks and firm-specific productivity shocks. Second, in Maćkowiak and Wiederholt (2009) we assume that the decision-maker chooses a law of motion for the signals and takes optimal actions given the signals. Here we assume that the decision-maker chooses directly a law of motion for the actions. Both setups yield the same behavior. See Section 6. Third, here the decision-maker chooses how much attention to devote to price setting. In Maćkowiak and Wiederholt (2009) the amount of attention devoted to price setting is fixed, and the decision-maker only chooses how to allocate this fixed amount of attention between aggregate conditions and idiosyncratic conditions.

In the economy there are multiple shocks. An agent with unlimited attention would respond perfectly to all shocks. Our agent with limited attention thinks of the problem in pieces. How should my price respond to aggregate technology? How should my price respond to monetary policy? How should my price respond to local conditions? Agents choose how much attention they devote to aggregate technology, monetary policy and local conditions. The more attention they devote to aggregate technology, the more closely p_{it}^{A} tracks its perfect information target p_{it}^{*A}. The more attention they devote to monetary policy, the more closely p_{it}^{R} tracks p_{it}^{*R}. The more attention they devote to local conditions, the more closely p_{it}^{I} tracks p_{it}^{*I}.

Turning to the labor mix decision, equation (24) characterizes the profit-maximizing labor mix in period t and equation (26) characterizes the labor mix chosen by the decision-maker. Here $\hat{w}_{jt}$ denotes the relative wage rate for type j labor in period t, $\tilde{\eta}$ is the wage elasticity of labor demand, and χ controls the amount of noise in the action. The information-processing error ν_{ijt}^{L} has the same properties as the information-processing errors in the equation for the price and is independent of all other information-processing errors.[17]

[17] We put more structure on the labor mix decision than on the price setting decision by expressing the labor mix

3.3 Loss in utility in the case of suboptimal actions

Let us turn to households. We now derive an expression for the loss in utility that a household incurs if the household chooses a consumption vector and a wage rate that deviate from the optimal decisions under perfect information. This expression is derived from the preferences presented in Section 2, the flow budget constraint presented in Section 2 and the demand function for type j labor.

We guess the following labor demand function

$$L_{jt} = \zeta \left(\frac{W_{jt}}{W_t} \right)^{-\tilde{\eta}} L_t. \tag{28}$$

We verify below that the equilibrium labor demand function has this form. Here L_{jt} is demand for type j labor in period t, L_t is the aggregate composite labor input, W_{jt} is the wage rate for type j labor, and $W_t = h\left(W_{1t}, \ldots, W_{Jt}\right)$ is a wage index. The function h is homogenous of degree one, continuously differentiable and symmetric. The coefficients $\tilde{\eta}$ and ζ satisfy $\tilde{\eta} > 1$ and $\zeta = J^{-\frac{\eta-\tilde{\eta}}{\eta-1}}$.

Substituting the consumption aggregator (2), the flow budget constraint (3) and the labor demand function (28) into the period utility function (1) yields the following expression for period utility

$$\frac{1}{1-\gamma} \left(\frac{R_{t-1} B_{jt-1} - B_{jt} + \left(1+\tau_w\right) W_{jt} \zeta \left(\frac{W_{jt}}{W_t} \right)^{-\tilde{\eta}} L_t + \frac{D_t}{J} - \frac{T_t}{J}}{\sum_{i=1}^{I-1} P_{it} \hat{C}_{ijt} + P_{It} \left(1 - \sum_{i=1}^{I-1} \hat{C}_{ijt}^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}} \right)^{1-\gamma}$$

$$-\frac{1}{1-\gamma} - \varphi \zeta \left(\frac{W_{jt}}{W_t} \right)^{-\tilde{\eta}} L_t,$$

where $\hat{C}_{ijt} = \left(C_{ijt}/C_{jt}\right)$ denotes relative consumption of good i. The denominator in the first term is consumption expenditure per unit of composite consumption. Dividing the numerator and the

as a function of relative wages rather than fundamental shocks. We do this because from equation (26) we derive the labor demand function and a labor demand function specifies labor demand on and off the equilibrium path. By expressing the labor mix as a function of relative wages, we specify labor demand on and off the equilibrium path.

denominator in the first term by the price index P_t yields

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

$$-\frac{1}{1-\gamma}-\varphi\zeta\left(\frac{\tilde{W}_{jt}}{\tilde{W}_t}\right)^{-\tilde{\eta}}L_t, \tag{29}$$

where $\Pi_t = (P_t/P_{t-1})$ is inflation, $\tilde{B}_{jt} = (B_{jt}/P_t)$ are real bond holdings, $\tilde{W}_{jt} = (W_{jt}/P_t)$ is the real wage rate for type j labor, $\tilde{D}_t = (D_t/P_t)$ are real profits, and $\tilde{T}_t = (T_t/P_t)$ are real lump-sum taxes.

Computing a log-quadratic approximation to the period utility function (29) around the non-stochastic steady state yields a simple expression for the loss in utility that a household incurs if the household chooses a consumption vector and a wage rate that deviate from the optimal decisions under perfect information. Let x_t denote the vector of variables appearing in the period utility function (29) that the household can affect in period t

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

Let $E_{j,-1}$ denote the expectation operator conditioned on the information of household j in period -1. After the log-quadratic approximation to the expected discounted sum of period utility, the expected discounted sum of losses in utility in the case of suboptimal actions equals

$$\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}\left(x_t-x_t^*\right)'H_0\left(x_t-x_t^*\right)+\left(x_t-x_t^*\right)'H_1\left(x_{t+1}-x_{t+1}^*\right)\right], \tag{30}$$

where the matrix H_0 and the matrix H_1 are given by

$$H_0 = -C_j^{1-\gamma}\begin{bmatrix} \gamma\omega_B^2\left(1+\frac{1}{\beta}\right) & \gamma\omega_B\tilde{\eta}\omega_W & 0 & \cdots & 0 \\ \gamma\omega_B\tilde{\eta}\omega_W & \tilde{\eta}\omega_W\left(\gamma\tilde{\eta}\omega_W+1\right) & 0 & \cdots & 0 \\ 0 & 0 & \frac{2}{\theta I} & \cdots & \frac{1}{\theta I} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \frac{1}{\theta I} & \cdots & \frac{2}{\theta I} \end{bmatrix}, \tag{31}$$

$$
H_1 = C_j^{1-\gamma} \begin{bmatrix} \gamma\omega_B^2 & \gamma\omega_B\tilde{\eta}\omega_W & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{32}
$$

and the vector of optimal decisions under perfect information x_t^* is given by

$$
\begin{aligned}
\tilde{b}_{jt}^* = \frac{1}{\beta}\left(r_{t-1} - \pi_t + \tilde{b}_{jt-1}^*\right) + \frac{\tilde{\eta}}{\tilde{\eta}-1}\frac{\omega_W}{\omega_B}\left[(1-\tilde{\eta})\,\tilde{w}_{jt}^* + \tilde{\eta}\tilde{w}_t + l_t\right] + \frac{\omega_D}{\omega_B}\tilde{d}_t - \frac{\omega_T}{\omega_B}\tilde{t}_t \\
-\frac{1}{\omega_B}c_{jt}^* - \frac{1}{\omega_B}\left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right),
\end{aligned} \tag{33}
$$

$$
c_{jt}^* = E_t\left[-\frac{1}{\gamma}\left(r_t - \pi_{t+1} - \frac{1}{I}\sum_{i=1}^{I}\left(\hat{p}_{it+1} - \hat{p}_{it}\right)\right) + c_{jt+1}^*\right], \tag{34}
$$

$$
\tilde{w}_{jt}^* = \gamma c_{jt}^* + \left(\frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right), \tag{35}
$$

$$
\hat{c}_{ijt}^* = -\theta\left(\hat{p}_{it} - \frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right). \tag{36}
$$

Here E_t denotes the expectation operator conditioned on the entire history of the economy up to and including period t, and the ω's in equations (31)-(33) are the following steady state ratios

$$
\begin{pmatrix} \omega_B & \omega_W & \omega_D & \omega_T \end{pmatrix} = \begin{pmatrix} \frac{\tilde{B}_j}{C_j} & \frac{\tilde{W}_j L_j}{C_j} & \frac{\frac{\tilde{D}}{J}}{C_j} & \frac{\frac{\tilde{T}}{J}}{C_j} \end{pmatrix}. \tag{37}
$$

In summary, 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 equations (33)-(36). Furthermore, for any process $\{x_t\}_{t=0}^{\infty}$, the expected discounted sum of losses in utility due to deviations of bond holdings, the wage rate and the consumption mix from the optimal decisions under perfect information is given by expression (30).[18] See Proposition 2 in Maćkowiak and Wiederholt (2010).

[18] To be precise, in the derivation of expression (30) we use a technical condition stated in Proposition 2 in Maćkowiak and Wiederholt (2010). This technical condition places some restrictions on the process $\{x_t\}_{t=0}^{\infty}$. This technical condition is satisfied in the paper.

3.4 Attention problem of a household

From the last subsection we know how much utility a household loses if the household takes actions that differ from the optimal actions under perfect information. We now state the attention problem of the household. We assume that the household allocates attention so as to maximize expected utility net of the cost of paying attention.

The attention problem of a household is similar to the attention problem of the decision-maker in a firm. There are two differences. First, the household chooses consumption but the variable appearing in objective (30) are real bond holdings. Therefore, we have to use the flow budget constraint to map consumption deviations into bond deviations. Second, the household makes two decisions that are closely related: the consumption decision and the wage setting decision. Since households have linear disutility of labor, the equation for the optimal real wage rate reads $\tilde{w}_{jt} = \gamma c_{jt}$. This intratemporal optimality condition states that the real wage rate should equal the marginal rate of substitution between consumption and leisure. For simplicity, we assume that households set real wage rates (not nominal wage rates). Then, households choose all variables in this intratemporal optimality condition, and since households know their own decisions, households satisfy this intratemporal optimality condition independent of information flows. This simplifies the households' attention problem.[19]

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

$$\max_{\kappa,B(L),C(L),\tilde{\theta},\xi} \left\{ \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[\frac{1}{2} \left(x_t - x_t^*\right)' H_0 \left(x_t - x_t^*\right) + \left(x_t - x_t^*\right)' H_1 \left(x_{t+1} - x_{t+1}^*\right) \right] - \frac{\lambda}{1-\beta}\kappa \right\}, \tag{38}$$

where

$$x_t - x_t^* = \begin{pmatrix} \tilde{b}_{jt} \\ \tilde{w}_{jt} \\ \hat{c}_{1jt} \\ \vdots \\ \hat{c}_{I-1jt} \end{pmatrix} - \begin{pmatrix} \tilde{b}_{jt}^* \\ \tilde{w}_{jt}^* \\ \hat{c}_{1jt}^* \\ \vdots \\ \hat{c}_{I-1jt}^* \end{pmatrix}, \tag{39}$$

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

$$\tilde{b}_{jt} - \tilde{b}_{jt}^* = -\sum_{l=0}^{t} \left(\frac{1}{\beta}\right)^l \left[\frac{1}{\omega_B} \left(c_{j,t-l} - c_{j,t-l}^*\right) + \tilde{\eta}\frac{\omega_W}{\omega_B} \left(\tilde{w}_{j,t-l} - \tilde{w}_{j,t-l}^*\right) \right], \tag{40}$$

[19] We also solved the model assuming that households set nominal wage rates. See Section 6.

the law of motion for the optimal actions under perfect information

$$c_{jt}^{*}=\underbrace{A_{1}\left(L\right)\varepsilon_{t}^{A}}_{c_{jt}^{*A}}+\underbrace{A_{2}\left(L\right)\varepsilon_{t}^{R}}_{c_{jt}^{*R}},\tag{41}$$

$$\tilde{w}_{jt}^{*}=\gamma c_{jt}^{*},\tag{42}$$

$$\hat{c}_{ijt}^{*}=-\theta\hat{p}_{it},\tag{43}$$

the law of motion for the actions

$$c_{jt}=\underbrace{B_{1}\left(L\right)\varepsilon_{t}^{A}+C_{1}\left(L\right)\nu_{jt}^{A}}_{c_{jt}^{A}}+\underbrace{B_{2}\left(L\right)\varepsilon_{t}^{R}+C_{2}\left(L\right)\nu_{jt}^{R}}_{c_{jt}^{R}},\tag{44}$$

$$\tilde{w}_{jt}=\gamma c_{jt},\tag{45}$$

$$\hat{c}_{ijt}=-\tilde{\theta}\left(\hat{p}_{it}+\frac{Var\left(\hat{p}_{it}\right)}{\xi}\nu_{ijt}^{I}\right),\tag{46}$$

and the constraint on the information content of the actions

$$\mathcal{I}\left(\left\{c_{jt}^{*A},c_{jt}^{*R},\hat{c}_{1jt}^{*},\ldots,\hat{c}_{I-1jt}^{*}\right\};\left\{c_{jt}^{A},c_{jt}^{R},\hat{c}_{1jt},\ldots,\hat{c}_{I-1jt}\right\}\right)\leq\kappa.\tag{47}$$

A household subject to rational inattention compares cost and benefit of paying attention. The benefit of paying attention to the current state of the economy is that actions get closer on average to the optimal actions under perfect information. The first term in objective (38) is the expected loss in utility if the law of motion for the actions differs from the law of motion for the optimal actions under perfect information. See the previous subsection for the derivation of this expression.

The first element of $x_t - x_t^*$ in equation (39) is the deviation of real bond holdings of the household from the real bond holdings that the same household would have if the household made the optimal decisions under perfect information in every period. In other words, $\tilde{b}_{jt} - \tilde{b}_{jt}^*$ is the deviation of real bond holdings of the household from real bond holdings of a hypothetical household with perfect information. The second element of $x_t - x_t^*$ is the deviation of the real wage rate of the household from the real wage rate of a hypothetical household with perfect information. The remaining elements of $x_t - x_t^*$ are the deviation of the consumption mix of the household from the consumption mix of a hypothetical household with perfect information. The bond deviation in period t is given by equation (40). This equation follows from the flow budget constraint (33) and says that the bond deviation in period t depends on current and past consumption deviations as well as current and past wage deviations.

Equations (41)-(43) characterize the optimal actions under perfect information, i.e., the actions of the hypothetical household with perfect information. We guess that optimal consumption under perfect information is a linear function of current and past shocks, a guess that we verify below. $A_1(L)$ and $A_2(L)$ are infinite-order lag polynomials. The optimal real wage rate under perfect information can be expressed as a function of optimal consumption under perfect information. The optimal consumption mix under perfect information can be expressed as a function of the relative prices of consumption goods.

The household chooses the law of motion for consumption and the real wage rate. See equations (44)-(46). $B_1(L)$ and $B_2(L)$ and $C_1(L)$ and $C_2(L)$ are infinite-order lag polynomials. For example, if the household chooses $B_1(L) = A_1(L)$ and $C_1(L) = 0$ and $B_2(L) = A_2(L)$ and $C_2(L) = 0$, the household decides to respond perfectly with consumption to aggregate shocks. The variables ν_{jt}^A, ν_{jt}^R, and ν_{ijt}^I are information-processing errors. These variables follow Gaussian white noise processes that are independent of fundamentals, independent across households, independent of each other, and have unit variance. Equation (45) is the standard intratemporal optimality condition stating that the real wage rate should equal the marginal rate of substitution between consumption and leisure. Since households have linear disutility of labor, this condition simply says that the real wage rate should equal a constant times consumption. Deviating from this optimality condition would imply utility losses, and the household can satisfy this condition independent of information flows because the household knows its own wage decision and consumption decision. Finally, by choosing the coefficients $\tilde{\theta}$ and ξ the household chooses the price elasticity of demand and the amount of noise in the consumption mix decision.[20]

The household would like to take actions that equal the optimal actions under perfect information in every period. If attention were free, this is what the household would do. However, we assume that paying attention is costly. Formally, the household faces constraint (47). The

[20] We put more structure on the consumption mix decision than on the intertemporal consumption decision by expressing the consumption mix as a function of relative prices rather than fundamental shocks. We do this because from equation (46) we derive the demand function and a demand function specifies demand on and off the equilibrium path. By expressing the consumption mix as a function of relative prices, we specify demand on and off the equilibrium path.

left-hand side of constraint (47) measures the amount of information acquired and processed by the household. The amount of information acquired and processed by the household is measured by the amount of information contained in the household's actions about the optimal actions under perfect information. Constraint (47) says that if the household wants to take more informed actions, the household has to pay more attention. Objective (38) states that paying attention is costly. The variable κ is the amount of attention devoted to the intertemporal consumption decision (how much to consume), wage setting, and the intratemporal consumption decision (which goods to consume). The parameter $\lambda > 0$ is the per-period marginal cost of paying attention.

3.5 Aggregation

In this subsection we describe how we aggregate decisions by individual firms and households. Log-linearizing equation (8) for aggregate output and the equations in (12) for aggregate composite consumption and aggregate composite labor input yields

$$y_t = \frac{1}{I} \sum_{i=1}^{I} \left(\hat{p}_{it} + y_{it} \right), \tag{48}$$

and

$$c_t = \frac{1}{J} \sum_{j=1}^{J} c_{jt}, \qquad l_t = \frac{1}{I} \sum_{i=1}^{I} l_{it}. \tag{49}$$

Log-linearizing the price index $P_t = d\left(P_{1t}, \ldots, P_{It}\right)$ and the wage index $W_t = h\left(W_{1t}, \ldots, W_{Jt}\right)$ yields

$$p_t = \frac{1}{I} \sum_{i=1}^{I} p_{it}, \qquad w_t = \frac{1}{J} \sum_{j=1}^{J} w_{jt}. \tag{50}$$

We also work with log-linearized equations when we aggregate the demand for good i and type j labor

$$c_{it} = \frac{1}{J} \sum_{j=1}^{J} c_{ijt}, \qquad l_{jt} = \frac{1}{I} \sum_{i=1}^{I} l_{ijt}. \tag{51}$$

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

$$y_{it} = a_t + a_{it} + \alpha l_{it}, \tag{52}$$

and

$$r_t = \rho_R r_{t-1} + \left(1 - \rho_R\right) \left[\phi_\pi \pi_t + \phi_y \left(y_t - y_t^P \right) \right] + \varepsilon_t^R. \tag{53}$$

4 Solving and evaluating the model

In this section we solve and evaluate the model.

4.1 Solving the model

We solve for the rational expectations equilibrium of the model using an iterative procedure. First, we make a guess concerning the law of motion for the profit-maximizing price, p_{it}^{*}, and a guess concerning the law of motion for the utility-maximizing consumption, c_{jt}^{*}. Second, we solve the firms' attention problem (21)-(27) and the households' attention problem (38)-(47). We turn each problem into a finite-dimensional problem by parameterizing each infinite-order lag polynomial $B(L)$ and $C(L)$ as a lag polynomial of a finite-order ARMA process.[21] Third, we aggregate the individual prices to obtain the price level. We aggregate across households to obtain 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 Taylor rule (53), $\pi_t = p_t - p_{t-1}$, and $y_t = c_t$. Finally, we compute the law of motion for the profit-maximizing price from equation (19) and the law of motion for the utility-maximizing consumption from equation (34). If the law of motion for the profit-maximizing price or the law of motion for the utility-maximizing consumption differs from our guess, we update the guess until a fixed point is reached.[22]

[21] We use Matlab and a standard nonlinear optimization program to solve the firms' problem and the households' problem. When solving the firms' problem, we assume that the expectation operator $E_{i,-1}$ in objective (21) is the unconditional expectation operator. This is the simplest assumption one can make and is also the assumption made in Sims (2003). When solving the households' problem, we make the following assumption concerning the expectation operator $E_{j,-1}$ in objective (38). We assume that households have perfect information up to and including period -1 and the particular realization of shocks up to and including period -1 is that shocks are zero. This assumption implies that all discounted second moments in objective (38) are finite even if $x_t - x_t^{*}$ has a unit root. We make this assumption because we want to allow for the possibility that $x_t - x_t^{*}$ has a unit root. Namely, we want to allow for the possibility that the bond holdings of a household subject to rational inattention differ asymptotically from the bond holdings of a hypothetical household with perfect information.

[22] On the way to the 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 computing time for a single iteration is at least one minute on a standard computer. The number of iterations needed to reach the fixed point depends significantly on parameter values, on the initial guess and on the weight of the guess in iteration $n-1$ in the guess in iteration n. Typically, we needed about 20-30 iterations to find the fixed point.

4.2 Evaluation strategy

We adhere to Thomas J. Sargent's dictum that "A rational expectations equilibrium *is* a likelihood function. Maximize it."[23] In an ideal world, we would choose values of the parameters of the model by likelihood-based estimation. We would evaluate the model by comparing its fit to the data with the fit of competing models, in terms of likelihood ratio or marginal likelihood. Smets and Wouters (2007) maximize the likelihood function, weighted by a prior, to select values of the parameters of a New Keynesian DSGE model. Smets and Wouters demonstrate that their model fits the data about as well, in terms of marginal likelihood, as a VAR. Ideally, we would like to show that the DSGE model with rational inattention achieves a similar value of marginal likelihood to the Smets-Wouters model and a VAR. The DSGE model with rational inattention holds promise in this regard, because this model relies on a single friction to explain the data and marginal likelihood favors a model with fewer parameters. However, we are not there yet. We cannot use likelihood-based estimation before advances in the speed of computation allow us to solve the model noticeably more quickly than we currently solve it. See Section 4.1.

Here we pursue the following strategy to set values of the parameters of the model. This strategy is computationally feasible and in the spirit of matching impulse responses as in Christiano et al. (2005) and Altig et al. (2011). We divide the parameters into two sets: the parameters specific to rational inattention (the marginal cost of information flow to the decision-maker in a firm, μ, and the marginal cost of information flow to a household, λ) and all other parameters (non-rational-inattention parameters). We calibrate the non-rational-inattention parameters based on extraneous information. We then solve the model for a grid of values of μ and λ. We select the pair of values of μ and λ that minimizes the distance between the impulse responses to a monetary policy shock in our model and the impulse responses to a monetary policy shock in the Smets-Wouters model. The details are in Section 4.3. We refer to the model with the parameter values chosen in this way as the *baseline economy*.[24]

[23] See Evans and Honkapohja (2005), p.567.

[24] We prefer to match the impulse responses in our model to the impulse responses in the Smets-Wouters model rather than in a VAR, for two reasons. First, there are multiple approaches to identify shocks in a VAR, each approach yields different results, and we do not want to rely on any single approach. Second, the Smets-Wouters model has become *the* benchmark model that economists use when they think of the effects of monetary policy in the business cycle. Therefore, we find it important to know whether our model matches the impulse responses to a monetary

Afterwards, we evaluate the model by comparing the impulse responses to a monetary policy shock and an aggregate technology shock in the baseline economy with the analogous impulse responses in the Smets-Wouters model, two other standard DSGE models and a standard VAR model. We also compare the unconditional moments in the baseline economy with the unconditional moments in the data. See Section 4.4.

4.3 Baseline economy: parameterization

Most of the non-rational-inattention parameters are standard and can be calibrated in a straightforward way.

One period in the model is one quarter. Therefore, we let $\beta = 0.99$. As is common in business cycle models, we set $\gamma = 1$ and $\alpha = 2/3$.

To calibrate the parameters of the stochastic process for aggregate technology and the parameters of the monetary policy rule, we use quarterly U.S. data from 1966 Q1 to 2004 Q4. We choose this sample period because Smets and Wouters (2007) use this sample period. As a measure of aggregate technology we employ total factor productivity, adjusted for utilization, reported by Fernald (2009).[25] We regress the log of TFP on a constant and a time trend. We then regress the residual on its own lag. Based on the point estimates from this regression, we set $\rho_A = 0.94$ and the standard deviation of ε_t^A equal to 0.008. These two numbers are typical to the literature on business cycles. We regress the Federal Funds rate on a constant, an own lag, a measure of the inflation rate and a measure of the output gap. We measure the output gap in the data as the difference between the log of real GDP and the log of real potential GDP reported by the Congressional Budget Office.[26] This is a standard empirical monetary policy rule. Based on the point estimates from this regression, we set $\rho_R = 0.9$, $\phi_\pi = 1.76$, $\phi_y = 0.4$ and a standard deviation of ε_t^R equal to 0.0022. These four numbers are typical to the literature on monetary policy rules.

We want monetary policy in the model to react to a similar measure of the output gap as in

policy shock in the Smets-Wouters model for some values of μ and λ.

[25] We downloaded the data from John Fernald's website, http://www.frbsf.org/economics/economists/staff.php?jfernald.

[26] As a measure of the inflation rate we use the difference in the log of the price index for personal consumption expenditures excluding food and energy. We downloaded the data on the Federal Funds rate, the price level, real GDP and real potential GDP from the website of the Federal Reserve Bank of St.Louis. We use the same data to compute the unconditional moments reported in the row "Data" in Table 3.

the empirical monetary policy rule. Typically, economists define the output gap in a model as the difference between output and frictionless output (here, output under perfect information). We find this approach unattractive in our context. Let y_t^{PI} denote output under perfect information. With $\gamma = 1$ we have $y_t^{PI} = a_t$. This is unappealing because in the data real potential GDP and TFP are very different series.[27] Therefore, we suppose that monetary policy in the model reacts to $y_t - y_t^P$ where y_t^P is smoothed output under perfect information, i.e., $y_t^P \equiv \zeta y_t^{PI} + (1-\zeta)\, y_{t-1}^P$, and we seek a realistic value of the parameter ζ. We compute a smoothed measure of the log of TFP in the data, $tfp_t^* \equiv \zeta tfp_t + (1-\zeta)\, tfp_{t-1}^*$. It turns out that for $\zeta = 0.05$ real potential GDP and the smoothed TFP are similarly behaved series.[28] Hence, we set $\zeta = 0.05$. In reality, central banks respond to a smooth measure of potential output and we make the same assumption in our model.

To set the parameters of the stochastic process for firm-specific productivity, we follow the literature calibrating menu cost models with firm-specific productivity shocks to U.S. micro price data. Nakamura and Steinsson (2008) and Klenow and Willis (2007) set the autocorrelation of firm-specific productivity equal to 0.66 and 0.68, respectively, at monthly rate. We set the autocorrelation of firm-specific productivity equal to 0.3, because our model is quarterly and $(0.3)^{1/3}$ equals a number between 0.66 and 0.68. Furthermore, Klenow and Kryvtsov (2008) report that the median absolute size of price changes (excluding sale-related price changes) equals 9.7 percent in the U.S. We set the standard deviation of the innovation to firm-specific productivity such that the median absolute size of price changes equals 9.7 percent in our model. This choice yields a standard deviation of ε_{it}^I equal to 0.23.[29]

When households and decision-makers in firms have perfect information, the price elasticity

[27] When modeled as a first-order autoregression, the growth rate of real potential GDP has an autocorrelation of 0.97 and a standard deviation of the innovation equal to 0.0009; by contrast, the growth rate of TFP has an autocorrelation of -0.12 and a standard deviation of the innovation equal to 0.0081.

[28] For $\zeta = 0.05$ the growth rate of tfp_t^* has an autocorrelation of 0.9 and a standard deviation of the innovation equal to 0.0009.

[29] To compute the median absolute size of price changes in our model we simulate prices in the baseline economy ($\mu = 0.0022$ and $\lambda = 0.002$, see below). An alternative approach would be to set the standard deviation of ε_{it}^I such that the median absolute size of price changes equals 9.7 percent in our model under perfect information ($\mu = \lambda = 0$). This alternative approach would yield the same standard deviation of ε_{it}^I. Note also that we match the average size of price changes excluding sale-related price changes. If we were to match the average size of all price changes, the standard deviation of ε_{it}^I would be larger.

of demand $\tilde{\theta}$ is equal to the preference parameter θ and the wage elasticity of labor demand $\tilde{\eta}$ is equal to the technology parameter η. When households and decision-makers in firms are subject to rational inattention, $\tilde{\theta}$ and $\tilde{\eta}$ are endogenous, $\tilde{\theta} < \theta$ and $\tilde{\eta} < \eta$. Households respond less strongly to fluctuations in relative prices of goods than under perfect information; and firms respond less strongly to fluctuations in relative wage rates than under perfect information. The modeler has the choice of *either* fixing θ and η and solving for $\tilde{\theta}$ and $\tilde{\eta}$ *or* fixing $\tilde{\theta}$ and $\tilde{\eta}$ and solving for θ and η. We interpret the empirical evidence on price elasticities of demand as coming from data generated by our model. Therefore, we fix $\tilde{\theta} = 4$ and $\tilde{\eta} = 4$. A price elasticity of demand of four is within the range of estimates of the price elasticity of demand in the Industrial Organization literature. When solving the model, we compute the parameter θ that yields $\tilde{\theta} = 4$ and the parameter η that yields $\tilde{\eta} = 4$. Furthermore, we set the number of consumption goods to $I = 100$ and the number of types of labor to $J = 100$. The parameter I has no effect on the responses of the household's composite consumption and the household's real wage rate to shocks and the parameter J has no effect on the responses of the firm's price to shocks. The parameter I only affects the parameter θ that yields $\tilde{\theta} = 4$ and J only affects the parameter η that yields $\tilde{\eta} = 4$.

We calibrate the parameter ω_B, the ratio of real bond holdings to consumption in the non-stochastic steady state, and the parameter ω_W, the ratio of real wage income to consumption in the non-stochastic steady state, based on data from the Survey of Consumer Finances. Our calibration strategy involves several steps and is described in Appendix B. The calibration strategy yields $\omega_B = 9.12$ and $\omega_W = 1.06$.

Holding the non-rational-inattention parameters constant at the selected values, we solve the model for a grid of values of the rational-inattention-parameters μ and λ. We find that the pair $\mu = 0.0022$ and $\lambda = 0.002$ minimizes the distance between the impulse responses of output and inflation to a monetary policy shock in our model and the impulse responses of the same variables to the same shock in the Smets-Wouters model.[30]

[30] For each pair of values of μ and λ on the grid, we calculate the sum of squared differences between the impulse responses of output (inflation) to a monetary policy shock in the two models, over the initial 20 quarters. For each pair k on the grid this yields two numbers: one for output, $\varsigma_k(y)$, and one for inflation, $\varsigma_k(\pi)$. We express each $\varsigma_k(y)$ as a percentage of the smallest $\varsigma_k(y)$ on the grid, i.e., for each k we compute $\varsigma_k(y) / \min(\varsigma(y))$ where $\min(\varsigma(y))$ is the smallest $\varsigma_k(y)$ on the grid. Analogously, we obtain $\varsigma_k(\pi) / \min(\varsigma(\pi))$ for each k. We then calculate $\varsigma_k(y) / \min(\varsigma(y)) + \varsigma_k(\pi) / \min(\varsigma(\pi))$ for each k. The pair $\mu = 0.0022$ and $\lambda = 0.002$ minimizes this function

We refer to the model with these parameter values as the baseline economy.

4.4 Baseline economy: results

We examine the properties of the baseline economy. The main takeaway is that the model fits the data about as well as standard DSGE models do; for this result, it is essential that the model include rational inattention on the side of firms *and* households.

We compare the baseline economy with two standard medium-sized DSGE models, a standard small DSGE model and data. By "two standard medium-sized DSGE models" we mean the Smets-Wouters model and the DSGE model in Altig et al. (2011), which we refer to as *ACEL DSGE*. Both models are popular benchmarks and are known to fit the data well. By "a standard small DSGE model" we mean the model that we refer to as the *Calvo-with-habit model*. The Calvo-with-habit model is a model with the same structure as the model in this paper except that: all agents have perfect information, decision-makers in firms face the Calvo friction in price setting, and households are subject to external habit formation in consumption. This is the most popular small DSGE model used to study the effects of monetary policy. We assume the same values for preference, technology and monetary policy parameters as in the baseline economy and we choose values for the Calvo parameter and the habit parameter in the same way as values of the rational-inattention-parameters before.[31] Finally, by a comparison with data we mean: (i) a comparison of impulse responses in the model and in the VAR model in Altig et al. (2011), which we refer to as *ACEL VAR*, and (ii) a comparison of unconditional moments in the model and in the data.

Figures 1-2 show the impulse responses of output, inflation and the nominal interest rate to a monetary policy shock (Figure 1) and an aggregate technology shock (Figure 2) in the baseline economy. The figures also display the analogous impulse responses in the Smets-Wouters model, the ACEL DSGE and the ACEL VAR.[32] Tables 1-2 show the standard deviation and the first-order

over k. The function $\varsigma_k(y)/\min(\varsigma(y)) + \varsigma_k(\pi)/\min(\varsigma(\pi))$ penalizes percentage deviations from the smallest sum of squared differences between the impulse responses of output (and inflation) in the two models. We repeat this procedure including the impulse responses of a third variable, the nominal interest rate, and obtain essentially the same result. In Section 6 we document the effects of varying μ and λ.

[31] We set the Calvo parameter equal to 0.81 and the habit parameter equal to 0.81 because these two values minimize the distance between the impulse responses of output and inflation to a monetary policy shock in the Calvo-with-habit model and the impulse responses of the same variables to the same shock in the Smets-Wouters model.

[32] All impulse responses are to shocks of one standard deviation. By an "aggregate technology shock" in Altig et

autocorrelation of output growth and inflation conditional on a monetary policy shock (Table 1) and conditional on an aggregate technology shock (Table 2) in the baseline economy, the Smets-Wouters model, the ACEL DSGE, the Calvo-with-habit model and the ACEL VAR.

The impulse response of output to a monetary policy shock in the baseline economy matches very well the Smets-Wouters model, the other DSGE models and the data. In particular, the path of output is hump-shaped with strongest response after about one year. The autocorrelation of output growth is between 0.5 and 0.6. See the top row of Figure 1 and Table 1.[33]

The impulse response of inflation to a monetary policy shock in the baseline economy matches less well the medium-sized DSGE models and the data. The size of the response in the baseline economy is correct and the response is about as persistent as in the medium-sized DSGE models and in the data. However, the impulse response of inflation in the baseline economy is monotonic, whereas the analogous impulse response is hump-shaped in the medium-sized DSGE models and in the data. See the middle row of Figure 1 and Table 1. Having said that, let us note the following. First, there is considerable uncertainty about when monetary policy's effect on inflation is strongest. The ACEL VAR suggests "after about a year and a half," but the Smets-Wouters model suggests only "after 2-3 quarters." The impulse response of inflation to a monetary policy shock in the baseline economy lies within the confidence interval for the analogous impulse response in the ACEL VAR. Second, the reason why both medium-sized DSGE models predict a hump-shaped impulse response of inflation to a monetary policy shock is that the medium-sized DSGE models assume an extra friction, price indexation. Third, inflation conditional on a monetary policy shock is more persistent in the baseline economy than in the Calvo-with-habit model. See Table 1. That is, rational attention produces more persistent inflation than does the combination of Calvo price setting and consumption habit.[34]

al. (2011) we mean their "neutral technology shock." The impulse responses in the Smets-Wouters model are shown with 80 percent posterior intervals. The impulse responses in the ACEL VAR are shown with 95 percent confidence intervals. Altig et al. (2011) assume that in any quarter pricing and consumption decisions are made before this quarter's monetary policy shock is realized. To facilitate comparison, we apply the same timing assumption in the baseline economy when we compare to the impulse responses in Altig et al. (2011) in the right column of Figure 1.

[33] The impulse response of output to a monetary policy shock is somewhat weaker in the ACEL DSGE and the ACEL VAR than in the other models. Note that the sample period in Altig et al. (2011) is 1982Q1-2008Q3. Recall that this paper and Smets and Wouters (2007) use the sample period 1966Q1-2004Q4.

[34] The impulse response of inflation to a monetary policy shock in the Calvo-with-habit model is monotonic.

The impulse response of output to an aggregate technology shock in the baseline economy matches well the other DSGE models and the data. In the baseline economy and the other DSGE models the response of output builds over time reaching its maximum after about two years. Output growth is autocorrelated.[35] In the ACEL VAR output is close to a random walk at the point estimate, but the confidence intervals are consistent with autocorrelation in output growth. See the top row of Figure 2 and Table 2.

The impulse response of inflation to an aggregate technology shock in the baseline economy is different than in the other DSGE models and matches the data better than these other DSGE models do. In the ACEL VAR inflation responds strongly on impact to an aggregate technology shock – an order of magnitude more strongly than to a monetary policy shock. The baseline economy matches this fact. By contrast, the other DSGE models fail to match this fact. In the other DSGE models the response of inflation to an aggregate technology shock is essentially the same as the response of inflation to a monetary policy shock. See the middle row of Figure 2 and Table 2.

Table 3 compares the unconditional moments in the model with the unconditional moments in the data. We consider three parameterizations of the model: the baseline economy, the model with rational inattention on the side of firms only (i.e., the baseline economy except that $\lambda = 0$) and the model with perfect information (i.e., the baseline economy except that $\mu = \lambda = 0$). As Table 3 shows, the data have two salient features: the autocorrelation of output growth is positive and the autocorrelation of inflation is close to one. The baseline economy matches the two features. Furthermore, rational inattention on the side of firms *and* on the side of households is essential for the model to match the two features. The model with perfect information matches neither feature. Under perfect information the autocorrelation of output growth is negative and the autocorrelation of inflation is close to zero. Rational inattention on the side of firms implies that the autocorrelation of inflation is close to one. The combination of rational inattention by firms *and* households implies that the autocorrelation of inflation is close to one *and* the autocorrelation of output growth is positive.

The autocorrelation of output growth is larger in the baseline economy than in the data (compare

[35] The response of output to an aggregate technology shock in the baseline economy is somewhat weaker than in the Smets-Wouters model.

0.68 with 0.24 in Table 3). It is important to note that in the model there are only two aggregate shocks, whereas presumably more than two shocks generate the variation in output in the data. If we added another aggregate shock to the model associated with a lower conditional autocorrelation of output growth, the unconditional autocorrelation of output growth would fall. To illustrate this point, consider the ACEL DSGE. The autocorrelation of output growth conditional on a monetary policy shock is the same in the ACEL DSGE and in the baseline economy. Furthermore, the autocorrelation of output growth conditional on an aggregate technology shock is almost the same in the ACEL DSGE and in the baseline economy. See Table 1 and Table 2, respectively. If the ACEL DSGE were a two-shock economy, the unconditional autocorrelation of output growth in that model would be 0.7, just like in the baseline economy. We think that in the future it will be worthwhile to add more shocks to our model to match not only the conditional moments quantitatively but also the unconditional moments quantitatively.[36]

Carroll et al. (2011) estimate that consumption growth has an autocorrelation of about 0.7. The model with rational inattention on the side of firms and households matches this evidence very well. In Table 3 note that the autocorrelation of output growth in the baseline economy is 0.68 and recall that consumption equals output in the model. Of course, the point from the previous paragraph applies here as well: Adding another aggregate shock may affect the unconditional autocorrelation of consumption growth predicted by the model.

We conclude the evaluation of the model with the following summary: The model fits the data about as well as the standard DSGE models do. For this result, it is essential that the model include rational inattention on the side of firms *and* households.

Next, we develop intuition for why the impulse responses in the baseline economy look the way they do. To this end, we examine the behavior of decision-makers in firms and households.

4.5 Understanding the behavior of firms in the DSGE model

The decision-maker in a firm faces the question: Given the marginal cost of paying attention μ, what is the optimal amount of attention allocated to monetary policy, aggregate technology and firm-specific productivity? We find that – of the total attention devoted to the price-setting decision – 4

[36] Adding another shock to the model would not change the impulse responses to the shocks already present in the model because of the constant marginal cost of paying attention.

percent is allocated to monetary policy, 13 percent to aggregate technology and 83 percent to firm-specific productivity. This allocation of attention implies that prices respond slowly to monetary policy shocks, fairly quickly to aggregate technology shocks and very quickly to market-specific shocks. Figure 3 illustrates this differential speed of response. The price of an individual firm converges to the firm's profit-maximizing price after 4 years following a monetary policy shock (top panel), after 2 years following an aggregate technology shock (second panel from the top) and after one quarter following a firm-specific productivity shock (third panel from the top).[37] The price shows dampened and delayed dynamics compared with the profit-maximizing price – as in a model with exogenous dispersed information – but the extent of dampening and delay depends on the optimal allocation of attention.

The endogenous attention allocation allows the model to: (i) match the empirical finding by Boivin et al. (2009) and Maćkowiak et al. (2009) that prices respond very quickly to disaggregate shocks, and (ii) produce a much stronger response of inflation to an aggregate technology shock than to a monetary policy shock, consistent with the ACEL VAR. Furthermore, the endogenous attention allocation implies that monetary policy has strong real effects *while profit losses are small*. In any model with a price setting friction, firms experience profit losses due to deviations of the price from the profit-maximizing price. In the baseline economy, the expected per-period loss in profit due to deviations of the price from the profit-maximizing price equals 0.24 percent of the firm's steady state revenue.[38] This number is *thirty-five times smaller* than the analogous number in the Calvo-with-habit model! The reason why profit losses are so much smaller in the rational inattention model is that in this model prices respond slowly only to unimportant shocks but quickly to important shocks.

Why is this attention allocation optimal? In the baseline economy the profit-maximizing price has the following property: little of its variation is due to monetary policy shocks, some of its variation is due to aggregate technology shocks and most of its variation is due to firm-specific

[37] The impulse response of the price level to each of the two aggregate shocks equals the impulse response of the price of an individual firm to that shock.

[38] The expected per-period profit loss due to imperfect tracking of monetary policy equals 0.02 percent of the firm's steady state revenue. The expected per-period profit loss due to imperfect tracking of aggregate technology equals 0.06 percent of the firm's steady state revenue. The expected per-period profit loss due to imperfect tracking of firm-specific productivity equals 0.16 percent of the firm's steady state revenue.

productivity shocks. Consider Figure 3. The absolute response of the profit-maximizing price to a monetary policy shock (top panel) is several times smaller than to an aggregate technology shock (second panel from the top). Furthermore, the absolute response of the profit-maximizing price to an aggregate technology shock is an order of magnitude smaller than to a firm-specific productivity shock (third panel from the top). This property of the profit-maximizing price incentivizes decision-makers in firms to absorb slowly information about monetary policy, fairly quickly information about aggregate technology and very quickly information about firm-specific productivity.

Why do the impulse responses of the profit-maximizing price differ so much by shock? One reason is that the non-rational-inattention parameters have been chosen to match key features of the data: the large average absolute size of price changes and the small standard deviation of the innovation in the monetary policy rule. The other reason is that there are *feedback effects*: the profit-maximizing price of a firm varies as decision-makers in other firms and households change their allocation of attention.

Let us explain the feedback effects by focusing on monetary policy shocks. To begin, suppose that only a single firm is subject to rational inattention, while all other firms and households have perfect information. This is the case with the feedback effects *switched off.* The profit-maximizing price response of this single firm to a monetary policy shock is shown in Figure 3 (bottom panel). The profit-maximizing price response is *several times larger* in absolute terms than in the baseline economy (top panel in Figure 3). Therefore, this single firm chooses to allocate *several times more* attention to monetary policy than in the baseline economy where all other firms and households also have limited attention. This allocation of attention implies that the firm's price converges to the profit-maximizing price after 6 quarters, compared with 4 years in the baseline economy.

Next, suppose that all other firms also become subject to rational inattention. The profit-maximizing price is given by

$$p_{it}^{*} = p_t + \frac{\frac{1-\alpha}{\alpha} + \gamma}{1 + \frac{1-\alpha}{\alpha}\tilde{\theta}} c_t - \frac{\frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\tilde{\theta}} a_t - \frac{\frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\tilde{\theta}} a_{it}. \tag{54}$$

When all firms absorb information slowly, the price level falls less after a positive innovation in the monetary policy rule; furthermore, consumption falls. These effects feed back into the profit-maximizing price response to a monetary policy shock via equation (54). With our parameterization, the profit-maximizing price moves *much less* and hence the optimal attention allocation shifts to *much less* attention to monetary policy.

Finally, suppose that all households also become subject to rational inattention. Additional feedback effects arise. The impulse response of consumption to a monetary policy shock becomes hump-shaped instead of monotonic (see Section 4.6) and thus the profit-maximizing price response to a monetary policy shock changes again. This induces another reallocation of attention by decision-makers in firms. The reallocation of attention can go either way, because two effects work in opposite directions: the profit-maximizing price moves less on impact (as consumption moves less on impact) but more after a few quarters (as the reaction of consumption is strongest then). It turns out that the optimal attention allocation by decision-makers in firms shifts to slightly *more* attention to monetary policy. While the reallocation of attention by firms is not large – because of the two effects working in opposite directions – the fact that rational inattention by households changes the profit-maximizing price is important. After we add rational inattention on the side of households, the impulse response of inflation to a monetary policy shock in the model matches much better the analogous object in the Smets-Wouters model.[39]

Feedback effects also appear in Maćkowiak and Wiederholt (2009) and in Hellwig and Veldkamp (2009). In this model feedback effects are richer than in that previous work. In the previous work there is one type of agent. Here there are two types of agents. The optimal attention allocation of a firm depends on the attention allocation of other firms and on the attention allocation of households.[40] Furthermore, the strength of the feedback effects differs across shocks. Feedback effects are stronger for monetary policy shocks than for aggregate technology shocks. That is, when we introduce rational inattention, the profit-maximizing price response to a monetary policy shock changes by more than the profit-maximizing price response to an aggregate technology shock. The reason is that the profit-maximizing price response to a monetary policy shock depends only on endogenous variables, p_t and c_t, whereas the profit-maximizing price response to an aggregate technology shock depends in addition on an exogenous variable, a_t. The effect of a_t on the profit-

[39] The reason is as follows. The impulse response of inflation to a monetary policy shock in the Smets-Wouters model is well approximated by an AR(2) process and is *not* well approximated by an AR(1) process. The same is true of the impulse response of inflation to a monetary policy shock in the baseline economy. By contrast, the impulse response of inflation to a monetary policy shock in the model with rational inattention on the side of firms only (i.e., the baseline economy except that $\lambda = 0$) is well approximated by an AR(1) process.

[40] Analogously, the optimal attention allocation of a household depends on the attention allocation of other households and on the attention allocation of firms.

maximizing price is present with constant strength no matter what attention choices agents make. See equation (54).[41] Finally, whether prices of different firms are strategic complements or strategic substitutes is more complicated to check than in the previous work. The reason is that the variables p_t and c_t appearing in equation (54) for the profit-maximizing price are no longer linked through a simple equation such as $c_t = m_t - p_t$, where m_t is an exogenous variable. Instead, c_t depends on p_t through the consumption Euler equation and the monetary policy rule.

4.6 Understanding the behavior of households in the DSGE model

Each household faces the question: Given the marginal cost of paying attention λ, what is the optimal amount of attention allocated to monetary policy and aggregate technology? The answer depends on the utility-maximizing consumption. The utility-maximizing consumption equals minus the sum of current and future real interest rates. In the baseline economy, it turns out that the real interest rate is driven to about the same extent by monetary policy shocks and aggregate technology shocks. Hence the household's problem in effect reduces to the choice of how much attention to pay to the real interest rate. We find that households decide to pay *little* attention to the real interest rate. This finding is important because in a large class of models monetary policy affects the economy *exclusively* via the real interest rate. Furthermore, this problem has not been studied in the literature on rational inattention before. Sims (2003, 2006), Luo (2008) and Tutino (2012) also study consumption-saving problems under rational inattention but assume that the real interest rate is constant.

To save space, let us focus on the household's consumption-saving decision conditional on a monetary policy shock.[42] The utility-maximizing response of consumption to a monetary policy shock in the baseline economy is depicted in Figure 4 (top panel). Since decision-makers in firms are rationally inattentive, the price level responds slowly to a monetary policy shock, the real interest rate rises after a positive innovation in the monetary policy rule and the utility-maximizing consumption falls on impact returning to zero monotonically.

Consider an individual household in the baseline economy. The impulse response of consumption of an individual household to a monetary policy shock is shown in Figure 4 (top panel). We find that

[41] The feedback effects are absent in the case of firm-specific productivity shocks, because the profit-maximizing price response to a firm-specific productivity shock depends only on an exogenous variable, a_{it}.

[42] The findings are the same conditional on an aggregate technology shock.

the household decides to allocate little attention to the real interest rate. Therefore, the impulse response of consumption is very different from the impulse response of the utility-maximizing consumption. The former is *hump-shaped*, whereas the latter is *monotonic*. When a contractionary monetary policy shock arrives, the household consumes too much and saves too little relative to what maximizes utility under perfect information. Afterwards, the household consumes persistently less – thereby slowly rebuilding its savings – compared with what would be the case if the household had perfect information.[43,44]

The difference between consumption and the utility-maximizing consumption is large and persistent, despite the fact that *utility losses are small!* In the baseline economy, the expected per-period loss in utility due to deviations of consumption and the real wage rate from the optimal decisions under perfect information equals the utility equivalent of 0.08 percent of the household's steady state consumption. In other words, to fully compensate the household for the expected discounted sum of utility losses due to deviations of consumption and the real wage rate from the optimal decisions under perfect information, it would be sufficient to give the household 1/1250 of the household's steady state consumption in every period. This result is important because in a large class of models the real interest rate is *the* transmission channel through which monetary policy affects the real economy. We find that paying little attention to the interest rate – and therefore responding slowly to changes in the real interest rate – is associated with small utility losses.

Next, think again about the utility-maximizing consumption. The utility-maximizing consumption depends on the non-rational-inattention parameters and via feedback effects on the optimal attention allocation by decision-makers in firms and households. Let us explain the feedback effects between households. Suppose that all firms and only a single household are subject to rational inattention, while all other households have perfect information. This is the case with the feedback effects between households *switched off*. The utility-maximizing response of consumption of this

[43] The impulse response of aggregate consumption to a monetary policy shock equals the impulse response of consumption of an individual household to that shock. The same is true for aggregate technology shocks.

[44] The difference between consumption and the utility-maximizing consumption goes to zero in the long run. Similarly, real bond holdings of the household differ from the utility-maximizing bond holdings but the difference between the two (not reported here) goes to zero in the long run. We also find that the impulse responses of consumption and real bond holdings under rational inattention to the noise terms in equation (44) go to zero in the long run. This finding implies that neither the cross-sectional variance of consumption nor the cross-sectional variance of real bond holdings diverges to infinity.

single household to a monetary policy shock is shown in Figure 4 (middle panel). The utility-maximizing response is *smaller* in absolute terms than in the baseline economy (Figure 4, top panel). Therefore, this single household chooses to allocate *less* attention to the real interest rate and responds *less* with its consumption to a monetary policy shock, compared with households in the baseline economy.

Suppose now that all households become subject to rational inattention. The feedback effects between households arise. The main insight is that the feedback effects between households are driven by the monetary policy rule, i.e., by the fact that monetary policy responds systematically to the state of the economy. When a contractionary monetary policy shock arrives, the real interest rate increases. The size of this increase is always attenuated because output decreases and output enters the Taylor rule. When all households absorb information slowly, output decreases *less* and therefore the real interest rate increases *more* and the utility-maximizing consumption decreases *more*. The optimal allocation of attention shifts to *more* attention to the real interest rate. Thus the feedback effects between households go in the *opposite* direction than the feedback effects between firms. If other households pay little attention to the real interest rate, the utility-maximizing consumption moves *more*, which raises the incentive for an individual household to attend to the real interest rate. Quantitatively, though, the feedback effects between households are of limited importance: Both the top panel and the middle panel of Figure 4 show large and persistent differences between consumption and utility-maximizing consumption.

In principle, it could be the case that a household subject to rational inattention pays little attention to the real interest rate only when the coefficient of relative risk aversion is low (or equivalently when the intertemporal elasticity of substitution is high). When the coefficient of relative risk aversion is low, deviations from the consumption Euler equation cost little in utility terms. Therefore, we also study what happens when we increase the coefficient of relative risk aversion by a factor of ten. We suppose again that all firms and only a single household are subject to rational inattention, while all other households have perfect information. In Figure 4 we compare the behavior of this single household when $\gamma = 1$ (middle panel) with the behavior of this single household when $\gamma = 10$ (bottom panel). As γ increases from 1 to 10, the attention devoted to the real interest rate rises by 50 percent and the ratio of the actual response to the utility-maximizing response of consumption on impact of a monetary policy shock doubles from 11 percent to 22 percent. The more risk-averse household devotes more attention to the real interest rate and therefore consumption responds faster to a monetary policy shock. However, the more risk-averse household still pays *little* attention to the real interest rate and its consumption still responds slowly to a monetary policy shock. This is because there are two effects working in opposite directions. When γ increases, so does the utility loss of a given deviation between consumption and the utility-maximizing consumption. See equation (30). This effect *raises* the attention devoted to the real interest rate. However, when γ increases, the coefficient on the real interest rate in the consumption Euler equation *falls*.[45] See equation (34). Thus the utility-maximizing consumption reacts *less* to a given change in the real interest rate. This effect *lowers* the attention devoted to the real interest rate. For γ between 1 and 10, the first effect dominates but only slightly.

In addition to the intertemporal consumption decision, the household also makes a decision about the consumption mix. The utility-maximizing consumption mix depends on the relative prices of goods. The relative prices of goods are driven by firm-specific productivity shocks and by mistakes in the firms' price-setting decisions. We find that households allocate *much more* attention to the consumption mix decision than to the intertemporal consumption decision. Specifically, the amount of attention households allocate to a *single* relative price of goods, i.e., $\hat{p}_{it}$ for a single i, equals about 70 percent of the attention allocated to the intertemporal consumption decision. Given that firm-specific productivity shocks induce large fluctuations in the relative prices of goods, households find it important to be aware of those fluctuations.

Let us restate the most important lessons of this subsection. Consumption responds slowly to changes in the real interest rate, because households decide to pay little attention to the real interest rate. This result: (i) holds for low and high degrees of relative risk aversion, and (ii) accounts for the model's fit to the data on output and consumption. Furthermore, feedback effects between households arise, driven by the monetary policy rule.

5 Experiments

In this section we use the model to conduct experiments.

To begin, we consider perhaps the most common experiment in the literature on DSGE models

[45] In other words, the intertemporal elasticity of substitution, $1/\gamma$, decreases.

used for analysis of monetary policy: we vary the coefficient on inflation in the monetary policy rule, ϕ_π. Figure 5 shows the effect of changing the value of ϕ_π on the standard deviation of the output gap due to aggregate technology shocks.[46] As the curve denoted "model with rational inattention" illustrates, in our model there is a *non-monotonic* relationship between ϕ_π and the volatility of the output gap.[47] In contrast, this relationship is *monotonic* in models with exogenous dispersed information and in simple New Keynesian models. Figure 5 illustrates this property for a version of our model in which we hold constant the amount of attention agents allocate to aggregate technology ("model with constant attention").

To understand how the value of ϕ_π affects the economy in the different models, note the following. As ϕ_π increases in models with exogenous dispersed information and in simple New Keynesian models, the nominal interest rate mimics more closely the real interest rate at the efficient solution (conditional on an aggregate technology shock). This standard effect *decreases* deviations of output from the efficient solution (again, conditional on an aggregate technology shock). In the model with rational inattention, there is an additional effect. When the central bank reacts more aggressively to inflation, the price level becomes more stable, implying that decision-makers in firms decide to pay *less* attention to aggregate conditions. This additional effect *increases* deviations of output from the efficient solution. If the second effect dominates, the volatility of the output gap increases. In Figure 5 the second effect dominates for empirically relevant values of ϕ_π, between 1.05 to 1.75. We conclude that the DSGE model with rational inattention gives a very different answer than the conventional DSGE models to the basic question what happens to the real economy when monetary policy fights inflation more aggressively.[48,49]

[46] In Section 5 we measure the output gap as the deviation of output from equilibrium output under perfect information. The latter equals efficient output, due to the subsidies (10)-(11).

[47] The relationship between ϕ_π and the standard deviation of the output gap is non-monotonic also conditional on monetary policy shocks.

[48] Paciello and Wiederholt (2012) show that optimal monetary policy in a model with rational inattention on the side of firms calls for complete price stability in response to aggregate technology shocks, i.e., it is optimal to set ϕ_π to infinity. Under the optimal policy, inflation and the output gap equal zero in every period. We show that as ϕ_π increases the real economy can become less, not more, stable when the economy is driven by aggregate technology shocks. In other words, the path to the optimal policy of Paciello and Wiederholt is non-monotonic. As ϕ_π increases, output gap volatility first increases and then decreases. It is optimal for the central bank to "go all the way" rather than strengthen its reaction to inflation incrementally.

[49] The value of ϕ_π also affects the cross-sectional distribution of prices in addition to the volatility of the output gap.

Next, we consider another common experiment in the literature on DSGE models used for analysis of monetary policy: we vary the degree of strategic complementarity in price setting. There is a large literature arguing that raising strategic complementarity in price setting *increases* real effects of monetary policy shocks. For example, Woodford (2003, Chapter 3) makes this point for New Keynesian models and Woodford (2002) makes this point for a model with exogenous dispersed information. Furthermore, it is common in the literature to argue that the degree of strategic complementarity in price setting is large and, hence, monetary policy shocks must have strong real effects. A well-known example is Altig et al. (2011) who argue that firms' marginal cost curves are significantly upward sloping in own output. Making firms' marginal cost curves more upward sloping in own output raises strategic complementarity in price setting and, in the model of Altig et al., strengthens real effects of monetary policy shocks.

Motivated by the literature, we consider the experiment of making firms' marginal cost curves more upward sloping in own output. In particular, we raise the degree of decreasing returns-to-scale, $1/\alpha$. As we decrease α from 1 to 0.5, real effects of monetary policy shocks first rise, peaking at 0.8, and then fall. There is a non-monotonic relationship between the degree of strategic complementarity in price setting and the standard deviation of the output gap due to monetary policy shocks. In Figure 6, compare the non-monotonic curve ("model with rational inattention") with the monotonic curve ("model with constant attention").

The reason for the non-monotonicity is that there are two effects. The first effect is the effect noted in the literature. With our parameterization, a decrease in α lowers the coefficient on consumption in the equation for the profit-maximizing price. See equation (54). In the language of Woodford (2003), a decrease in α raises the degree of strategic complementarity in price setting. In the language of Ball and Romer (1990), a decrease in α raises the degree of real rigidity. 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 equation (18) increases in absolute value. Decision-makers in firms therefore 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

A higher ϕ_π lowers the response of the profit-maximizing price to aggregate shocks. As a result, the tracking problem of price setters becomes easier, price setters make smaller idiosyncratic mistakes, and inefficient price dispersion falls.

shocks. In Figure 6 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 0.8. Hence, for reasonable parameter values, raising strategic complementarity in price setting *reduces* real effects. We conclude that the conventional wisdom that strategic complementarity in price setting necessarily strengthens real effects of monetary policy does not hold in the rational inattention DSGE model.[50,51]

Let us also describe what happens when we raise the standard deviation of any aggregate shock. Decision-makers in firms and households decide to pay *more* attention to the aggregate economy. This result explains the evidence in Coibion and Gorodnichenko (2012) who study survey data on expectations finding that the degree of attention to the aggregate economy rose markedly during the turbulent 1970s and fell significantly during the subsequent calm period. Furthermore, this result has implications for the effects of monetary policy. As the standard deviation of monetary policy shocks increases in the 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. The reallocation of attention is important quantitatively. For example, if we start in the baseline economy and double the standard deviation of monetary policy shocks, real effects of a monetary policy shock last 2 years instead of 4 years.

The lesson of this section is that the outcomes of experiments are very different in this model than in the DSGE models currently used to analyze monetary policy, even though the models yield similar impulse responses. See Section 4.4. Therefore, it matters which model one uses for monetary policy analysis.

[50] Reis (2006b) studies pricing of a firm that faces a fixed cost of acquiring, absorbing and processing information. He obtains a related result: When the profit function peaks more sharply, the firm acquires information more frequently.

[51] An increase in the price elasticity of demand, $\tilde{\theta}$, has analogous effects in the model to an increase in the degree of decreasing returns-to-scale, $1/\alpha$. It lowers the coefficient on consumption in equation (54) and raises in absolute value the upper-left element of the matrix H in equation (18). However, in principle one can raise strategic complementarity in price setting without changing the upper-left element of the matrix H. This is the case when there exists a parameter in the model that affects the former but not the latter. See Ball and Romer (1990) or more recently Nakamura and Steinsson (2010).

6 Extensions

In this section we document how the solution of the model changes when we vary assumptions concerning information flows, in three different ways.

To begin, we study the effects of changing the values of the rational-inattention parameters μ and λ from the baseline economy values $\mu = 0.0022$ and $\lambda = 0.002$. Recall from Section 4.3 that the baseline economy values of the parameters μ and λ minimize the distance between the impulse responses of output and inflation to a monetary policy shock in our model and the impulse responses of the same variables to the same shock in the Smets-Wouters model. Figure 7 shows how our model's fit to the Smets-Wouters model varies with μ and λ. The figure displays the criterion function we used in Section 4.3 to pick the best values of μ and λ. The criterion function is normalized so that at its minimum, for $\mu = 0.0022$ and $\lambda = 0.002$, the function attains the value of 100.[52] Thus, the figure shows by how much in percentage terms the fit deteriorates as one moves away from the minimum. We draw three conclusions. First, the procedure we used to select the values of μ and λ works well in the sense that the criterion function is clearly minimized at $\mu = 0.0022$ and $\lambda = 0.002$. We conclude that it is feasible to pick values of rational-inattention parameters in the same way as one picks values of any other parameter in a macroeconomic model, e.g., by matching impulse responses like here or by likelihood-based estimation. Second, the criterion function rises only moderately for small departures of μ from the best value and small departures of λ from the best value.[53] We conclude that the model is robust to small variation in the values of μ and λ in the sense that such variation deteriorates the model's fit only moderately. Third, non-trivial amounts of rational inattention on the side of firms *and* households are necessary for good fit. In Figure 7 start from $\mu = 0.0022$ and $\lambda = 0.002$, decrease *either* μ *or* λ by about one-half, and notice that the criterion function rises sharply.

Next, we state the attention problem of the decision-maker in a firm using signals. So far we have assumed that the decision-maker in a firm chooses directly a law of motion for the actions, subject to a constraint on the information content of the actions. We now assume that the decision-maker

[52] The definition of the criterion function is in Footnote 30.

[53] In Figure 7 note that the criterion function rises moderately for μ between 0.002 and 0.0026, so long as λ stays near 0.002; and that the criterion function rises moderately for λ between 0.0016 and 0.0028, so long as μ stays near 0.0022.

in a firm chooses the precision of signals in period -1, subject to a constraint on the information content of the signals. In the following periods, the decision-maker in a firm receives the signals and takes the optimal actions given the signal realizations.

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

$$\max_{(\kappa,\sigma_1,\sigma_2,\sigma_3,\sigma_4)\in\mathbf{R}^5_{++}} \left\{\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}\left(x_t-x_t^*\right)' H \left(x_t-x_t^*\right)\right]-\frac{\mu}{1-\beta}\kappa\right\}, \tag{55}$$

where

$$x_t - x_t^* = \begin{pmatrix} p_{it} \\ \hat{l}_{i1t} \\ \vdots \\ \hat{l}_{i(J-1)t} \end{pmatrix} - \begin{pmatrix} p_{it}^* \\ \hat{l}_{i1t}^* \\ \vdots \\ \hat{l}_{i(J-1)t}^* \end{pmatrix}, \tag{56}$$

subject to equations (23)-(24) characterizing the profit-maximizing actions in period t, the following equation characterizing the agent's actions in period t

$$x_t = E\left[x_t^* | \mathcal{F}_{i0}, s_{i1}, s_{i2}, \ldots, s_{it}\right], \tag{57}$$

the following equation characterizing the agent's signal vector in period t

$$s_{it} = \begin{pmatrix} p_{it}^{*A} \\ p_{it}^{*R} \\ p_{it}^{*I} \\ \hat{w}_{1t} \\ \vdots \\ \hat{w}_{(J-1)t} \end{pmatrix} + \begin{pmatrix} \sigma_1 \nu_{it}^A \\ \sigma_2 \nu_{it}^R \\ \sigma_3 \nu_{it}^I \\ \sigma_4 \nu_{i1t}^L \\ \vdots \\ \sigma_4 \nu_{i(J-1)t}^L \end{pmatrix}, \tag{58}$$

and the constraint on information flow

$$\mathcal{I}\left(\left\{p_{it}^{*A}, p_{it}^{*R}, p_{it}^{*I}, \hat{l}_{i1t}^*, \ldots, \hat{l}_{i(J-1)t}^*\right\}; \{s_{it}\}\right) \leq \kappa. \tag{59}$$

The random variables ν_{it}^A, ν_{it}^R, ν_{it}^I, and ν_{ijt}^L are signal noise. These variables follow Gaussian white noise processes that are independent of fundamentals, independent across firms, independent of each other, and have unit variance. $E_{i,-1}$ in objective (55) is the expectation operator conditioned on the information of the decision-maker in firm i in period -1. $\mathcal{F}_{i0}$ in equation (57) 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 signals in period -1) the agent receives information such that the conditional covariance matrix of x_t^* given information in period t is constant for all $t \geq 0$. The operator $\mathcal{I}$ in the information flow constraint (59) measures the information content of the signals. Finally, the parameter $\mu \geq 0$ in objective (55) is the marginal cost of paying attention.

We solve the problem (55)-(59) for an individual firm assuming that aggregate variables and relative wage rates are given by the equilibrium of the model presented in Section 4.4. We then compare the solution of problem (55)-(59) to the solution of problem (21)-(27). We find that the two solutions are *identical*. The impulse responses of the price set by the firm to the three fundamental shocks are identical in the two problems. The impulse responses of the price set by the firm to the noise terms in equation (25) and to the noise terms in equation (58) are also identical. This is a numerical result. We find it remarkable that signals with noise that is i.i.d. across time yield the same price setting behavior as the more flexible decision problem (21)-(27).[54]

Finally, we relax the assumption that paying attention to aggregate technology, paying attention to monetary policy, and paying attention to firm-specific productivity are independent activities. We replace the signal vector (58) by the following signal vector

$$s_{it} = \begin{pmatrix} p_t \\ a_t + a_{it} \\ c_{i,t-1} \\ w_{t-1} + l_{i,t-1} \\ \hat{w}_{1t} \\ \vdots \\ \hat{w}_{(J-1)t} \end{pmatrix} + \begin{pmatrix} \sigma_1 \nu_{i1t} \\ \sigma_2 \nu_{i2t} \\ \sigma_3 \nu_{i3t} \\ \sigma_4 \nu_{i4t} \\ \sigma_5 \nu_{i1t}^L \\ \vdots \\ \sigma_5 \nu_{i(J-1)t}^L \end{pmatrix}. \tag{60}$$

By choosing σ_1 to σ_5 the decision-maker decides how much attention to devote to the price level, total factor productivity, last period sales, the last period wage bill, and the relative wage rates.[55,56]

[54] See Maćkowiak and Wiederholt (2011) for the details of how we solve problem (55)-(59) and for an extended description of the results. Furthermore, note that in Maćkowiak and Wiederholt (2009), Propositions 3-4, we prove analytically in a special case that signals with noise that is i.i.d. across time yield the same price setting behavior as the more flexible decision problem (21)-(27).

[55] We maintain the assumptions about signal noise stated above.

[56] We include last period sales and last period wage bill in the signal vector because we do not know how the firm

The variables in the signal vector (60) are driven by multiple shocks and it is therefore no longer the case that attending to aggregate technology, attending to monetary policy, and attending to firm-specific productivity are independent activities. We find that solving the problem (55)-(59) with the signal vector (60) instead of the signal vector (58) changes the firm's price setting behavior hardly at all. The reason is that the decision-maker in the firm decides to pay close 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 (60) and obtained similar results. See Maćkowiak and Wiederholt (2011) for an extended description of the results.[57,58]

7 Conclusion

Making good decisions in an environment with shocks requires attention. We solve a DSGE model in which paying attention is costly and agents allocate attention optimally. The model fits macroeconomic data about as well as standard DSGE models do; but without Calvo price setting, without habit formation in consumption, without Calvo wage setting, and without price indexation. Having decision-makers in firms that pay little attention to monetary policy and households that pay little attention to the real interest rate is enough. Furthermore, this allocation of attention is optimal for most of postwar U.S. history. However, if the structure of the economy changes, e.g., if times become more turbulent, the optimal attention allocation will change and thus the propagation of shocks will change.

---

can attend to current period sales and current period wage bill before setting the price.

[57] 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 (60) is chosen optimally subject to the constraint on information flow (59). 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.

[58] As another extension, we 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. 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 consumption-saving decision.

A Quantifying information flows

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 constraints (27) and (47) is defined as

$$\mathcal{I}(\{X_t\}; \{Y_t\}) = \lim_{T \to \infty} \frac{1}{T} \left[H(X_0, \ldots, X_{T-1}) - H(X_0, \ldots, X_{T-1} | Y_0, \ldots, Y_{T-1}) \right], \tag{61}$$

where $\{X_t\}_{t=0}^{\infty}$ and $\{Y_t\}_{t=0}^{\infty}$ are stochastic processes. The operator $\mathcal{I}$ quantifies the information that one process contains about another process 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

$$\mathcal{I}(\{X_t\}; \{Y_t\}) = \lim_{T \to \infty} \frac{1}{T} \left[\frac{1}{2} \log_2 \left(\frac{\det \Omega_X}{\det \Omega_{X|Y}} \right) \right]. \tag{62}$$

If X_t is a scalar, Ω_X is the covariance matrix of the vector $(X_0, \ldots, X_{T-1})$. If X_t is itself a vector, Ω_X is the covariance matrix of the vector obtained by stacking the vectors $X_0, \ldots, X_{T-1}$. In practice, we evaluate (62) for a large but finite T. If a variable appearing in the information flow constraint is non-stationary, we replace the variable by its first difference to ensure that entropy is always finite.

B Calibration details

To calibrate the parameters ω_B and ω_W we use data from the Survey of Consumer Finances 2007 and proceed as follows. 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 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 "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 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_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 ω_W equal to the ratio of quarterly nominal wage income to our proxy for quarterly consumption expenditure: $\omega_W = (10283.75/9695.5) = 1.06$.

References

[1] Altig, David, Lawrence J. Christiano, Martin Eichenbaum and Jesper Lindé (2011): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle." Review of Economic Dynamics, 14(2), 225-247.

[2] Angeletos, George-Marios, and Jennifer La'O (2009a): "Incomplete Information, Higher-Order Beliefs and Price Inertia." Journal of Monetary Economics, 56(S), 19-37.

[3] Angeletos, George-Marios, and Jennifer La'O (2009b): "Noisy Business Cycles." NBER Macroeconomics Annual 2009, 319-378.

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

[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 (2011): "International Evidence on Sticky Consumption Growth." Review of Economics and Statistics, 93(4), 1135-1145.

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

[8] Coibion, Olivier, and Yuriy Gorodnichenko (2012): "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." Discussion paper, College of William and Mary and U.C. Berkeley.

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

[10] Evans, George W., and Seppo Honkapohja (2005): "An Interview With Thomas J. Sargent." Macroeconomic Dynamics, 9(4), 561–583.

[11] Fernald, John (2009): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity." Unpublished manuscript, Federal Reserve Bank of San Francisco.

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

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

[14] Justiniano, Alejandro, and Giorgio Primiceri (2008): "The Time Varying Volatility of Macroeconomic Fluctuations." American Economic Review, 98(3), 604-641.

[15] Kacperczyk, Marcin, Stijn Van Nieuwerburgh and Laura Veldkamp (2012): "Rational Attention Allocation over the Business Cycle." Discussion paper, New York University.

[16] 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.

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

[18] Leeper, Eric M. (1991): "Equilibria under 'Active' and 'Passive' Monetary and Fiscal Policies." Journal of Monetary Economics, 27(1), 129-147.

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

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

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

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

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

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

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

[26] Maćkowiak, Bartosz, and Mirko Wiederholt (2011): "Business Cycle Dynamics under Rational Inattention." ECB Working Paper 1331.

[27] 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.

[28] Matejka, Filip (2011): "Rationally Inattentive Seller: Sales and Discrete Pricing." Discussion paper, CERGE-EI.

[29] Matejka, Filip, and Christopher A. Sims (2010): "Discrete Actions in Information-Constrained Tracking Problems." Discussion paper, CERGE-EI and Princeton University.

[30] Mondria, Jordi (2010): "Portfolio Choice, Attention Allocation, and Price Comovement." Journal of Economic Theory, 145(5), 1837-1864.

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

[32] Nakamura, Emi, and Jón Steinsson (2010): "Monetary Non-Neutrality in a Multisector Menu Cost Model." Quarterly Journal of Economics, 125(3), 961-1013.

[33] Paciello, Luigi (2012): "Monetary Policy and Price Responsiveness to Aggregate Shocks under Rational Inattention." Journal of Money, Credit and Banking, forthcoming.

[34] Paciello, Luigi, and Mirko Wiederholt (2012): "Exogenous Information, Endogenous Information and Optimal Monetary Policy." Discussion paper, Einaudi Institute for Economics and Finance and Goethe University Frankfurt.

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

[36] Reis, Ricardo (2006b): "Inattentive Producers." Review of Economic Studies, 73(3), 793-821.

[37] Sims, Christopher A. (1998): "Stickiness." Carnegie-Rochester Conference Series on Public Policy, 49(1), 317-356.

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

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

[40] Sims, Christopher A. (2010): "Rational Inattention and Monetary Economics." In "Handbook of Monetary Economics", Volume 3, edited by Benjamin M. Friedman and Michael Woodford, Elsevier.

[41] 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.

[42] Tutino, Antonella (2012): "Rationally Inattentive Consumption Choices." Review of Economic Dynamics, forthcoming.

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

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

[45] Veldkamp, Laura L. (2011): Information Choice in Macroeconomics and Finance. Princeton University Press, Princeton.

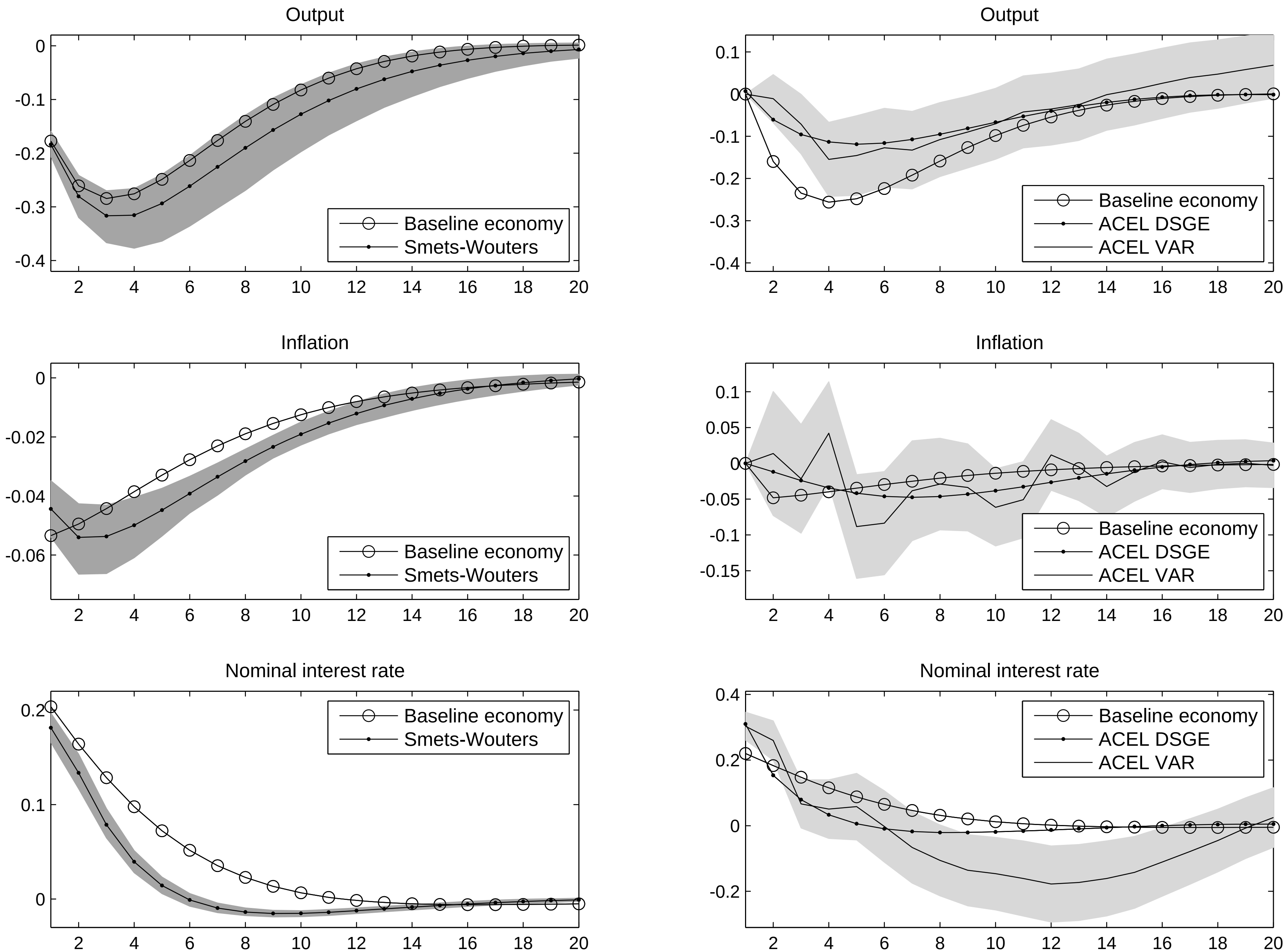
[46] Wiederholt, Mirko (2010): "Rational Inattention." In "The New Palgrave Dictionary of Economics", online edition, edited by Steven N. Durlauf and Lawrence E. Blume, Palgrave Macmillan.

[47] 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", edited by Philippe Aghion et al., Princeton and Oxford: Princeton University Press.

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

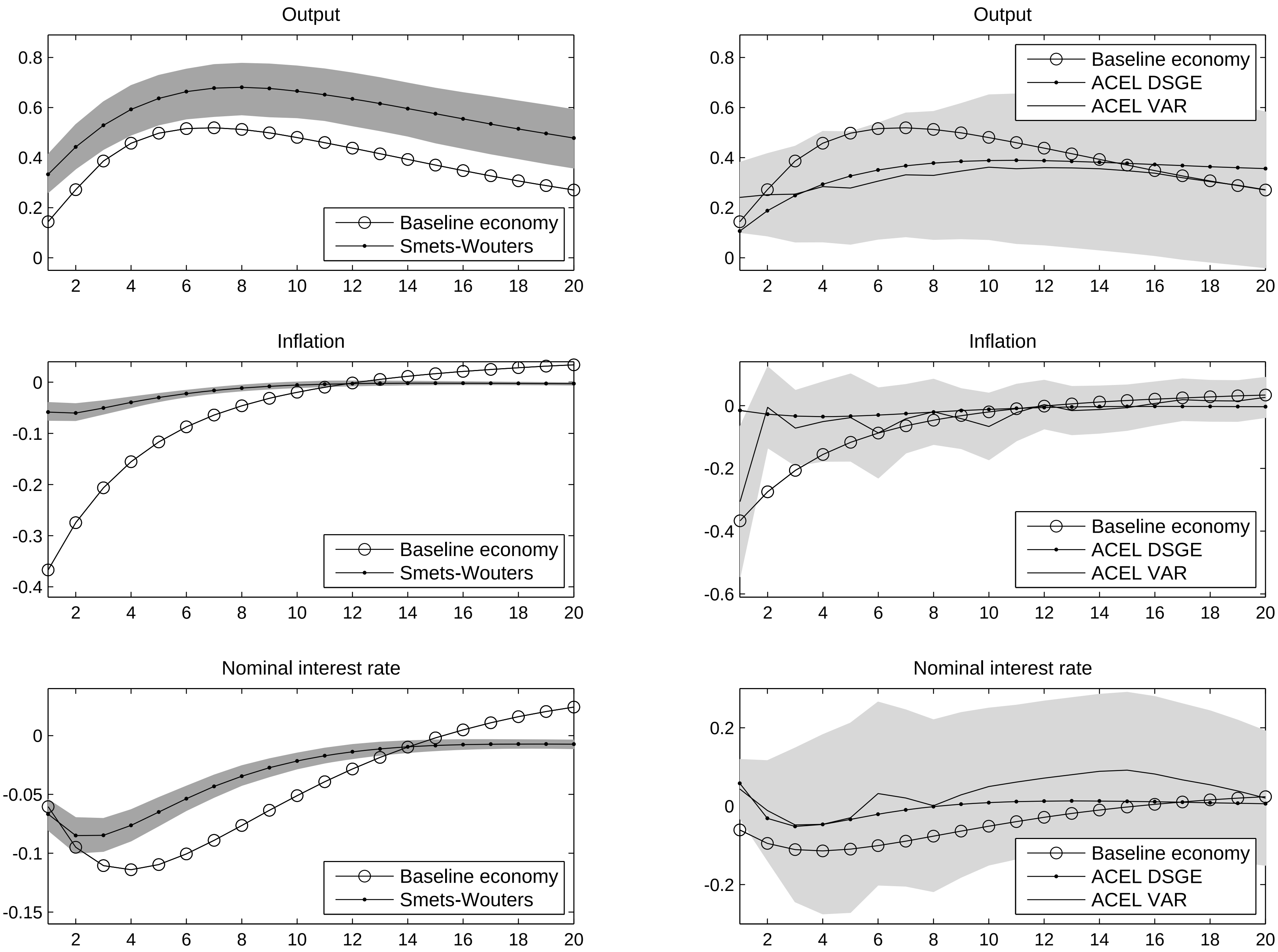
[49] Woodford, Michael (2009): "Information-Constrained State-Dependent Pricing." Journal of Monetary Economics, 56(S), 100-124.

Figure 1: Impulse responses to a monetary policy shock



Sources: Altig et al. (2011), Smets and Wouters (2007), and own calculations.
An impulse response equal to 1 means a 1 percent, or percentage point, deviation from the non-stochastic steady state. Time is measured in quarters.
The Smets-Wouters impulse responses are shown with 80 percent posterior intervals and the ACEL VAR impulse responses - with 95 percent confidence intervals.

Figure 2: Impulse responses to an aggregate technology shock



Sources: Altig et al. (2011), Smets and Wouters (2007), and own calculations.
An impulse response equal to 1 means a 1 percent, or percentage point, deviation from the non-stochastic steady state. Time is measured in quarters.
The Smets-Wouters impulse responses are shown with 80 percent posterior intervals and the ACEL VAR impulse responses - with 95 percent confidence intervals.

Table 1: Second moments conditional on a monetary policy shock

	Output growth		Inflation	
	Standard deviation	Autocorrelation	Standard deviation	Autocorrelation
Baseline economy	0.0022	0.51	0.0011	0.87
Smets-Wouters	0.0023	0.55	0.0013	0.93
ACEL DSGE	0.0009	0.51	0.0013	0.98
Calvo-with-habit	0.0023	0.58	0.0009	0.73
ACEL VAR	0.0012	0.52	0.0017	0.45

Sources: Altig et al. (2011), Smets and Wouters (2007) and own calculations.

Table 2: Second moments conditional on an aggregate technology shock

	Output growth		Inflation	
	Standard deviation	Autocorrelation	Standard deviation	Autocorrelation
Baseline economy	0.0025	0.80	0.0056	0.75
Smets-Wouters	0.0038	0.42	0.0011	0.85
ACEL DSGE	0.0016	0.75	0.0008	0.96
Calvo-with-habit	0.0011	0.86	0.0012	0.80
ACEL VAR	0.0025	0.07	0.0035	0.18

Sources: Altig et al. (2011), Smets and Wouters (2007) and own calculations.

Table 3: Unconditional second moments

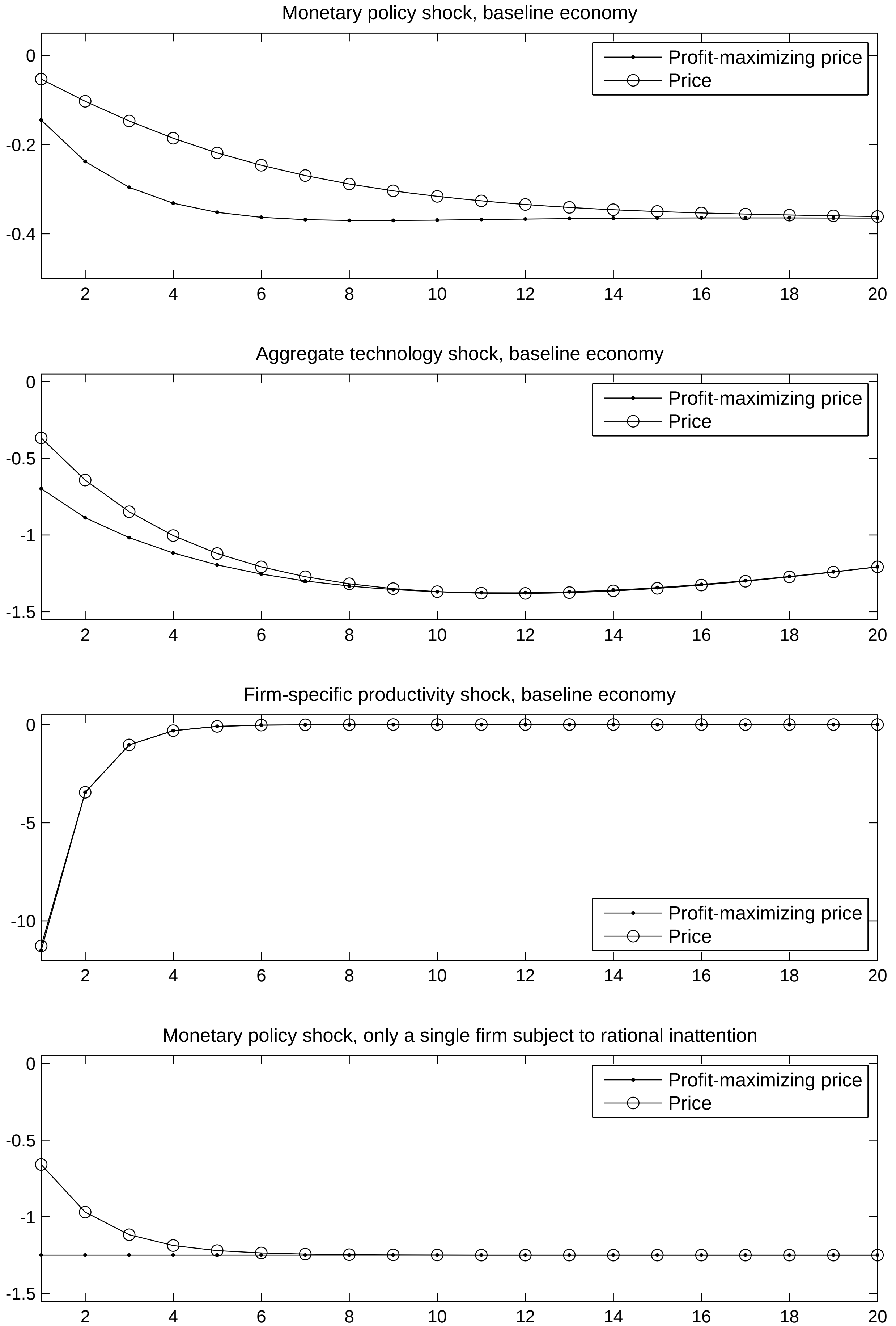
	Output growth		Inflation	
	Standard deviation	Autocorrelation	Standard deviation	Autocorrelation
Data	0.0084	0.24	0.0055	0.91
Baseline economy	0.0033	0.68	0.0057	0.75
Perfect information	0.0081	-0.03	0.0173	0.07
Only firms subject to RI	0.0090	-0.07	0.0038	0.81

Sources: Federal Reserve Bank of St. Louis and own calculations.

"Perfect information" is the same as "Baseline economy" except that $\mu=\lambda=0$.

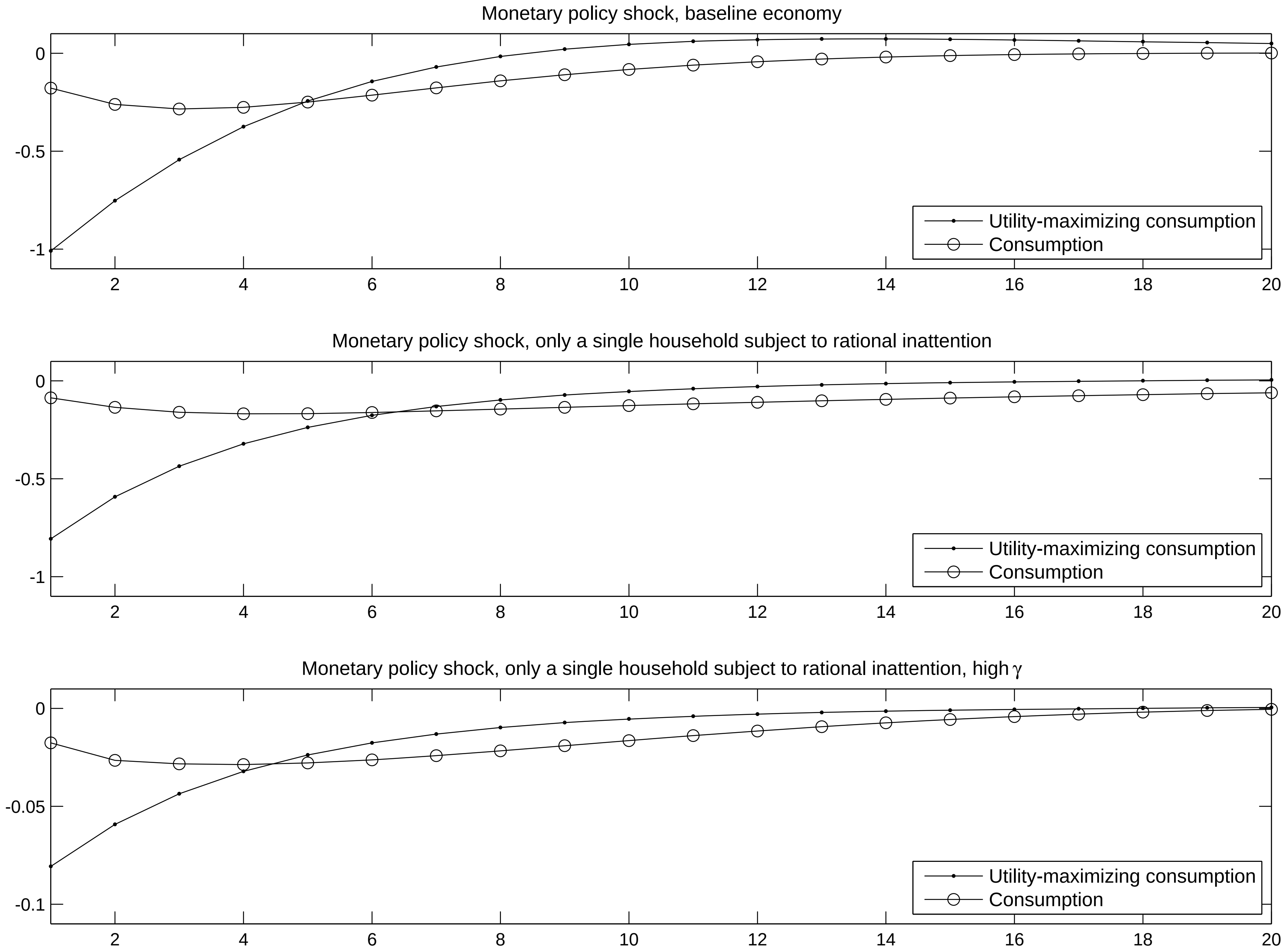
"Only firms subject to RI", where RI stands for rational inattention, is the same as "Baseline economy" except that $\lambda=0$.

Figure 3: Impulse responses of prices



Source: Own calculations. An impulse response equal to 1 means a 1 percent deviation from the non-stochastic steady state. Time is measured in quarters.

Figure 4: Impulse responses of consumption



Source: Own calculations. An impulse response equal to 1 means a 1 percent deviation from the non-stochastic steady state.
Time is measured in quarters.

Figure 5: Standard deviation of output gap due to aggregate technology shocks

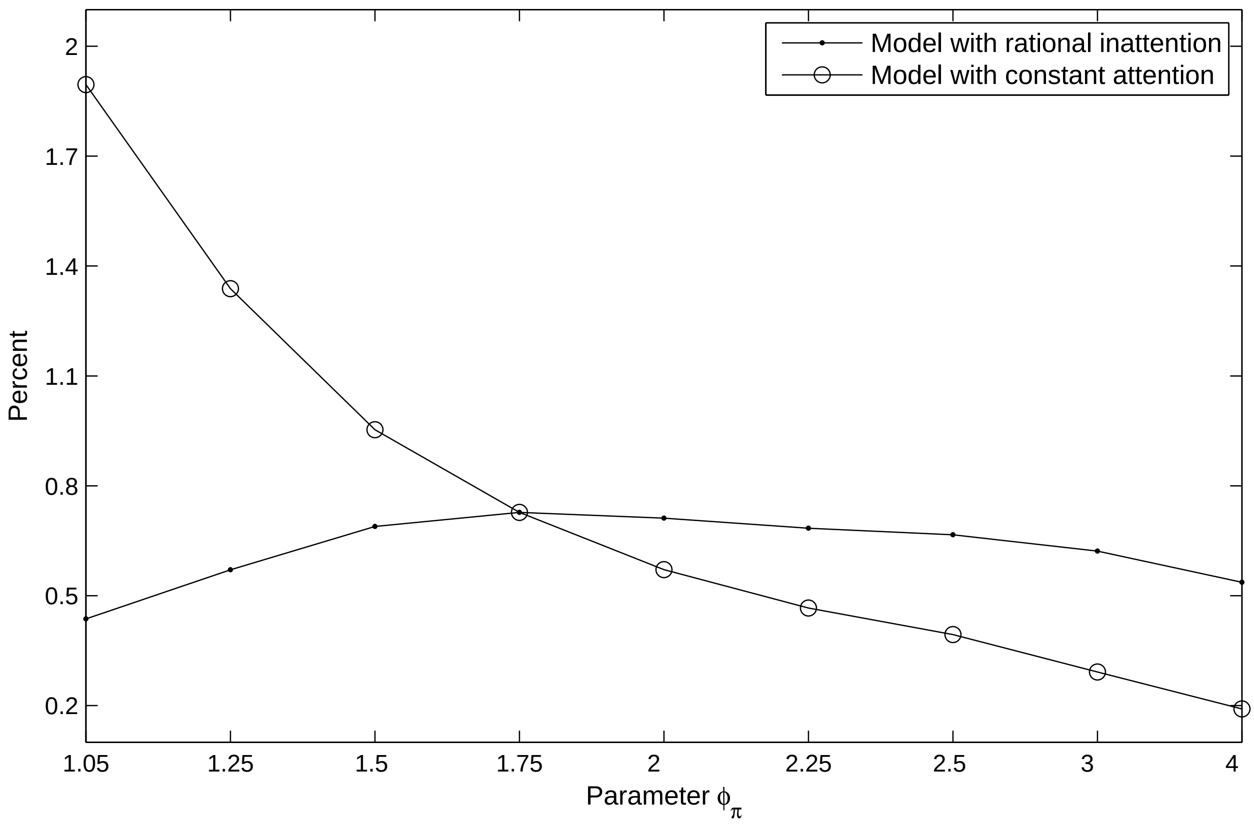
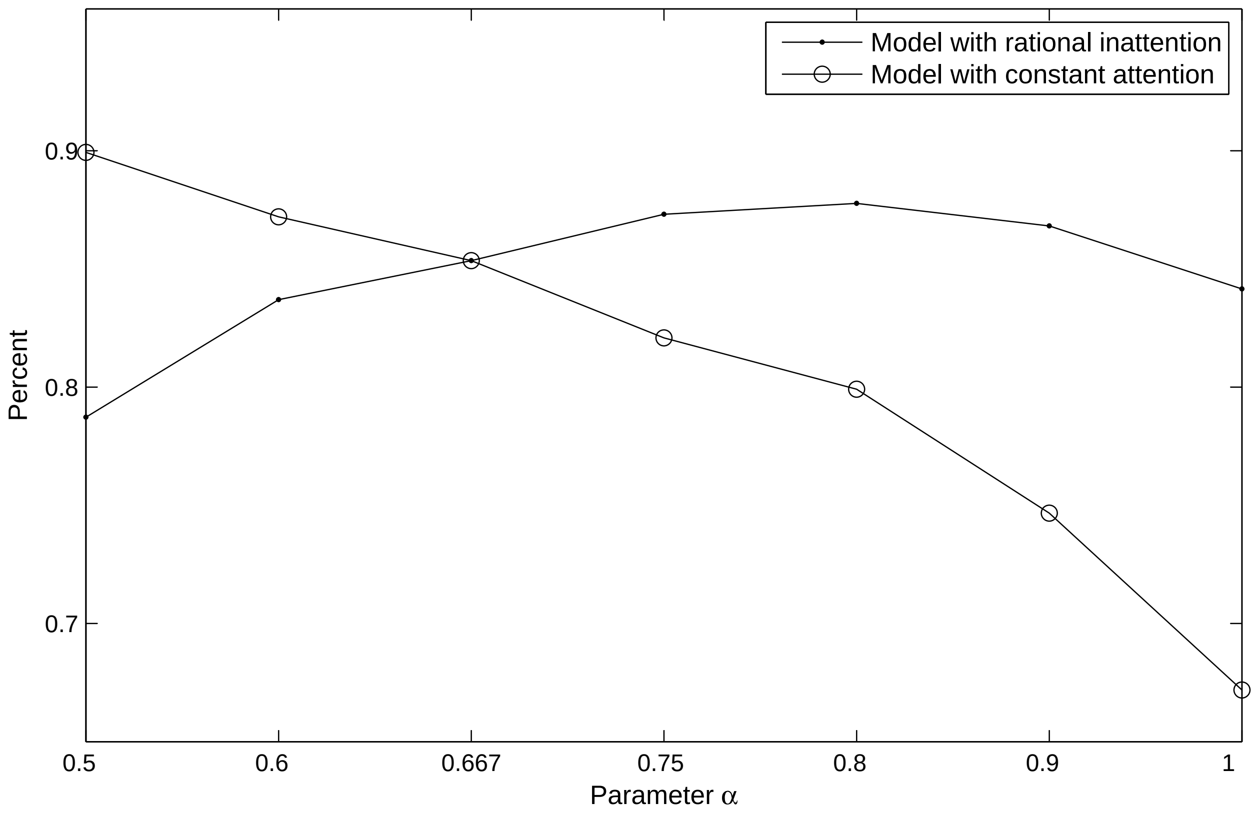
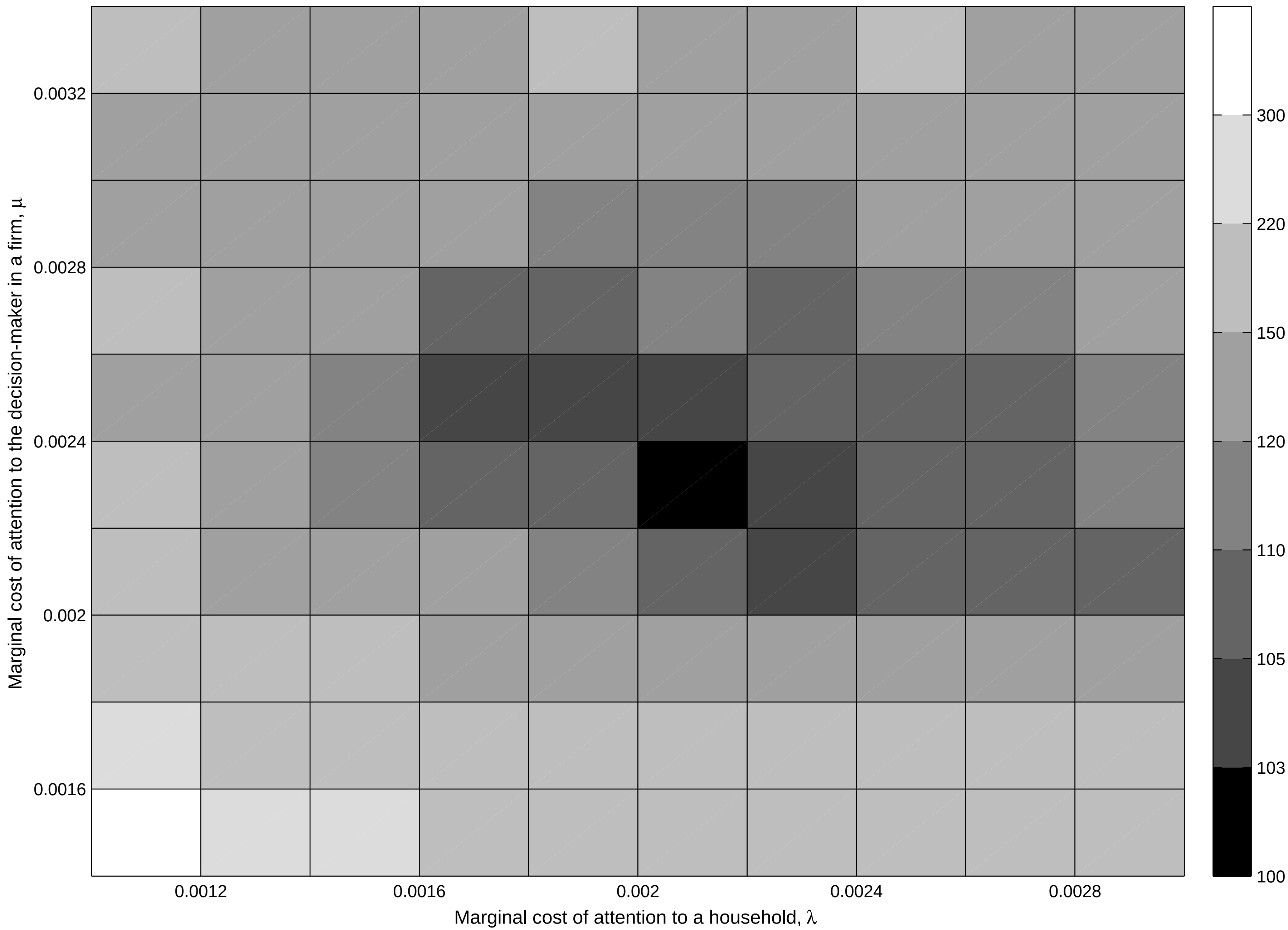


Figure 6: Standard deviation of output gap due to monetary policy shocks



Source for Figures 5-6: Own calculations. Figures 5-6 have on the vertical axes the standard deviation multiplied by 100, i.e., the number 1 means a standard deviation of 1 percent. See Section 5 for the details.

Figure 7: Fit to the Smets-Wouters model as a function of the parameters μ and λ



Source: Own calculations. Figure 7 shows the criterion function used in Section 4.3 to choose the best values of μ and λ. The criterion function is defined in Footnote 30. The criterion function is normalized so that at its minimum, for μ=0.0022 and λ=0.002, the function attains the value of 100.