Abstract

We develop a dynamic stochastic general equilibrium (DSGE) model with rational inattention and compare its predictions to data. Households and decision-makers in firms have limited attention and optimally allocate their attention. Rational inattention is the only source of slow adjustment. The model matches the empirical impulse responses to monetary policy shocks and aggregate technology shocks. At the same time, profit losses and utility losses from inattention are very small. Furthermore, it matters whether one uses this model or a conventional DSGE model for policy analysis.

Keywords: information choice, rational inattention, monetary policy, business cycles. (JEL: D83, E31, E32, E52).
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.\(^1\) Sims 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 develops such a model and compares it to data.

The model is close to a simple New Keynesian model, except that we discard all sources of slow adjustment that usually are in New Keynesian models (Calvo pricing, habit formation in consumption, Calvo wage setting) and we abandon the assumption of costless attention. There are many firms, many households, and a government. Firms supply differentiated goods that are produced 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. The central bank sets the nominal interest rate according to a Taylor rule. The economy is affected by aggregate technology shocks, monetary policy shocks, and firm-specific productivity shocks. Decision-makers in firms and households must pay attention in order to be aware of economic shocks. Decision-makers in firms and households choose how much attention to pay to the economic shocks, facing a constant marginal cost of attention. Following Sims (2003), attention is quantified as uncertainty reduction, where uncertainty is measured by entropy. After calibrating the preference, technology, and policy parameters, we select the marginal cost of attention for the decision-maker in a firm and the marginal cost of attention for a household so as to minimize the distance between the impulse responses to a monetary policy shock in the model and in a standard structural vector autoregression. The DSGE literature has developed by matching the impulse responses from VARs.

\(^1\)Later Christiano et al. (2005), Smets and Wouters (2007), Altig et al. (2011) and many others confirmed Sims’s conclusion in more formal analysis.
We follow the same approach.

The main findings are as follows. For the selected values of the marginal costs of attention, the model matches well the empirical impulse responses to a monetary policy shock. Furthermore, for the same values of the marginal costs of attention, the model also matches well the empirical impulse responses to an aggregate technology shock.

Let us give some details. Empirically, output follows a hump-shaped path after a monetary policy shock. The model reproduces this feature of the data. When everyone’s attention is costless, a monetary policy shock has no real effects. When firms’ attention is costly, prices evolve slowly after a monetary policy shock and therefore the shock affects output. When firms’ attention and households’ attention are costly, the impulse response of output to a monetary policy shock becomes hump-shaped rather than monotonic, because households adjust their consumption only gradually. Thus, rational inattention by firms and rational inattention by households are essential. Furthermore, for specific values of the two marginal costs of attention, the model matches the empirical impulse responses of inflation and output to a monetary policy shock not only qualitatively, but also quantitatively.

Let us turn to aggregate technology shocks. One important feature of the empirical impulse responses is that the price level responds much faster to aggregate technology shocks than to monetary policy shocks. For the selected values of the marginal costs of attention, the model matches this feature qualitatively and quantitatively. The model matches this feature qualitatively, because in the calibration we chose to match key features of postwar U.S. data and aggregate technology shocks appear to be much larger than monetary policy shocks in postwar U.S. data. Decision-makers in firms therefore decide to pay more attention to aggregate technology shocks than to monetary policy shocks, and as a result, prices respond faster to aggregate technology shocks than to monetary policy shocks. Surprisingly, the model matches the faster response of the price level to aggregate technology shocks not only qualitatively, but also quantitatively.

Under rational inattention actions deviate from optimal actions. (If a decision-maker processed perfectly all available information, the decision-maker would take different actions.) Since actions deviate from optimal actions, decision-makers incur a loss in utility or profit. We find that utility and profit losses from inattention are small, even though the model matches the inertia in the macroeconomic data and inattention is the only source of slow adjustment. Let us explain, starting
with households. Suppose that a contractionary monetary policy shock occurs, a household changes consumption by -0.1 percent, while the optimal consumption adjustment is -0.6 percent, i.e., the household consumes 0.5 percent “too much.” In the next quarter, the household offsets the effect of this consumption deviation on the present-value budget constraint by consuming the required amount less than dictated by the optimal consumption plan. With log utility, this simple double deviation from the optimal consumption path carries a utility loss of about \((0.005)^2 = (1/40,000)\) of steady-state consumption. Importantly, this is the size of a deviation that is sufficient to match the data. The only difference between this example and the consumption response to a one standard deviation monetary policy shock in the model is that households consume too much for seven quarters (which increases the utility loss) and spread out the adjustment that restores the present-value budget constraint over many quarters (which reduces the utility loss). These two effects cancel. The utility loss caused by a suboptimal consumption response to a one standard deviation monetary policy shock equals about \((1/40,000)\) of steady-state consumption. In addition, a household also incurs utility losses due to suboptimal labor market decisions and due to noise in the actions. Overall, the expected per-period utility loss due to suboptimal tracking of monetary policy shocks equals about \((1/5,900)\) of steady-state consumption. Since this utility loss is so small, a household is willing to tolerate this loss (i.e., chooses not to pay more attention to monetary policy) even if the marginal cost of attention is small.

Let us turn to firms. In the model decision-makers in firms decide to allocate a lot of attention to firm-specific productivity, some attention to aggregate technology, and little attention to monetary policy. Since this attention allocation is optimal, prices in the model respond very quickly to firm-specific productivity shocks, fairly quickly to aggregate technology shocks, and slowly to monetary policy shocks. This pattern matches the empirical evidence on the speed of response of prices to different shocks. Furthermore, due to this attention allocation, the profit loss from inattention is small and, at the same time, real effects of monetary policy shocks are large. Prices in the model respond slowly only to unimportant shocks (monetary policy shocks) but quickly to important shocks (firm-specific productivity shocks, followed by aggregate technology shocks). In our baseline parameterization, the expected per-period profit loss due to suboptimal price setting equals about \((1/1,700)\) of a firm’s steady-state revenue. This number is ninety times smaller than the analogous number in a Calvo model that produces similar real effects of monetary policy shocks.
Hence, the slow adjustment in macroeconomic data that is usually modelled with various forms of adjustment costs may actually have a different source: rational inattention.

We examine how changes in parameter values affect the incentives to pay attention. Let us focus on one result. We assume constant relative risk aversion preferences. When one increases the coefficient of relative risk aversion, households pay even less attention to the macroeconomy. There are two effects when the coefficient of relative risk aversion increases. The response of optimal consumption to a shock that moves the real interest rate is scaled down (which decreases the incentive to pay attention), and a given deviation between actual consumption and optimal consumption becomes more costly (which increases the incentive to pay attention). The first effect dominates.

We use the model to conduct experiments. We find that the outcomes of experiments are very different than in a New Keynesian model and in a model with exogenous dispersed information. The reason is intuitive. When a parameter value changes, decision-makers in this model reallocate attention. The reallocation of attention turns out to have a critical effect on the outcomes of experiments.

Consider a data-driven experiment. As is well known, the volatility of output growth and inflation declined in the 1980s. A popular hypothesis is that “the Great Moderation” was caused by a change in policy. We capture the shift to anti-inflationary monetary policy by the Federal Reserve under the chairmanship of Paul Volcker as an increase in the coefficient on inflation in the Taylor rule. When we raise this coefficient in the model from 1.01 to 1.5, the variance of output growth conditional on aggregate technology shocks declines and the variance of inflation conditional on aggregate technology shocks declines. By contrast, in a basic New Keynesian model, the variance of output growth conditional on aggregate technology shocks increases.

The intuition is the following. Let us focus on aggregate technology shocks. There are two effects when the coefficient on inflation in the Taylor rule increases. The first effect is that the nominal interest rate mimics more closely the real interest rate that implements the efficient output level in a world where households have perfect information. This effect decreases deviations of output from efficient output, i.e., the volatility of the output gap falls and the volatility of output rises. The second effect is that the price level becomes more stable, implying that decision-makers in firms decide to pay less attention to the macroeconomy. This effect increases deviations of output from
efficient output, i.e., the volatility of the output gap rises and the volatility of output falls. In our experiment, the second effect dominates and thus the volatility of output falls.

The second effect is absent from a New Keynesian model and from a model with exogenous dispersed information. Hence, in a basic New Keynesian model increasing the coefficient on inflation in the Taylor rule from 1.01 to 1.5 raises the volatility of output due to aggregate technology shocks. For this reason, change-in-policy explanations of the Great Moderation in the New Keynesian literature typically involve moving from an indeterminacy region to a determinacy region of the parameter space or involve other shocks. See, for example, Lubik and Schorfheide (2004).

Survey data on expectations provide direct evidence of the reallocation of attention predicted by the model. Coibion and Gorodnichenko (2012) study survey data on expectations finding that the degree of attention to the aggregate economy fell during the Great Moderation period, consistent with our model.

This paper belongs to the literature on rational inattention following Sims (2003).2 The main 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. 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.3

The optimal consumption response to a macroeconomic shock in the model depends on current and future real interest rates. Therefore, another way to express our result that households decide to pay little attention to macroeconomic shocks is to say that households choose to pay little

---


3 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.
attention to the real interest rate. Sims (2003, 2006), Luo (2008), and Tutino (2013) also study consumption-saving decisions under rational inattention but assume that the real interest rate is constant.\footnote{The real interest rate is constant also in Reis (2006).} Therefore, the point that households have little incentive to attend to the real interest rate (for high and low values of the intertemporal elasticity of substitution) 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), Mankiw and Reis (2002), Woodford (2002), Lorenzoni (2009), and Angeletos and La’O (2013)). 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.

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 describes the Great Moderation experiment. Section 6 considers extensions. Section 7 concludes. An online Appendix contains additional material.

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).\footnote{For reasons explained below, we make two additional changes to the standard New Keynesian model: asset markets are incomplete and there is a large finite number of firms (instead of a continuum of firms).} 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 sum of period utility. The discount factor is $\beta \in (0,1)$. The period utility function is

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

with

$$C_{jt} = \left( \sum_{i=1}^{I} C_{ijt}^{\theta} \right)^{\frac{1}{\theta}}.$$  \hspace{1cm} (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.\textsuperscript{6} The flow budget constraint of household $j$ in period $t$ reads

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

\textsuperscript{6}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.
Here $P_{it}$ is the price of good $i$ in period $t$, $R_{t-1}$ is the gross nominal interest rate on bond holdings between periods $t-1$ and $t$, $B_{jt-1}$ are bond holdings by household $j$ between periods $t-1$ and $t$, $W_{jt}$ is the nominal wage rate for labor supplied by household $j$ in period $t$, $\tau_\nu$ is a wage subsidy, $(D_t/J)$ is a pro-rata share of nominal aggregate profits, and $(T_t/J)$ is a pro-rata share of nominal lump-sum taxes. 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, $B_{jt} > 0$.

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.

### 2.2 Firms

There are $I$ firms. Firm $i$ supplies good $i$. The production function of firm $i$ is

$$Y_{it} = e^{\alpha_t} e^{\alpha_t} L_{it}^\alpha,$$

with

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

Here $Y_{it}$ is output, $e^{\alpha_t} e^{\alpha_t}$ is total factor productivity, $L_{it}$ is composite labor input, and $L_{ijt}$ is input of type $j$ labor at firm $i$ in period $t$. Type $j$ labor is the labor supplied by household $j$. Total factor productivity has an aggregate component, $e^{\alpha_t}$, and a firm-specific component, $e^{\alpha_t}$. 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

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

---

7 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.

8 We assume that there is a finite number of firms, because a household subject to rational inattention cannot track a continuum of prices. Dixit and Stiglitz (1977) also assume that there is a finite number of firms and that firms take the aggregate price index as given.
where $\tau_p$ is a production subsidy.

In every period, each firm sets a price, $P_{it}$, and chooses a labor mix, $(\hat{L}_{iti}, \ldots, \hat{L}_{i(J-1)t})$, where $\hat{L}_{ijt} = (L_{ijt}/L_{it})$ denotes relative input of type $j$ labor at firm $i$ 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 wage rates, the aggregate price index defined below and all aggregate quantities.

### 2.3 Government

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

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

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 equilibrium output under perfect information, and $\varepsilon_t^R$ is a monetary policy shock. The price index $P_t$ is defined later. $R$ and $\Pi$ denote the values of the nominal interest rate and inflation in the non-stochastic steady state. The policy parameters satisfy $\rho \in (0, 1)$, $\phi_\pi > 1$ 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 so as to correct the distortion arising from firms’ market power in the goods market, and the government sets the wage subsidy so as to correct the distortion arising from households’ market power in the labor market. Formally,

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

9
where \( \tilde{\theta} \) denotes the price elasticity of demand, and

\[
\tau_w = \frac{\tilde{\eta}}{\tilde{\eta} - 1} - 1,
\]

(11)

where \( \tilde{\eta} \) denotes the wage elasticity of labor demand.\(^9\)

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 \( \{\varepsilon_t^R\}, \{a_t\}, \{a_{1t}, ..., a_{It}\} \) are independent. The variable \( \varepsilon_t^R \) follows a Gaussian white noise process, \( a_t \) follows a stationary Gaussian first-order autoregressive process with mean zero, and each \( a_{it} \) 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 \( \varepsilon_t^A \) and \( \varepsilon_{it}^I \), respectively.

### 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

\[
C_t = \sum_{j=1}^{J} C_{jt}, \quad L_t = \sum_{i=1}^{I} L_{it}, \quad (12)
\]

\( \hat{P}_{it} \) denotes the relative price of good \( i \) and \( \hat{W}_{jt} \) denotes the relative wage rate for type \( j \) labor

\[
\hat{P}_{it} = \frac{P_{it}}{P_t}, \quad \hat{W}_{jt} = \frac{W_{jt}}{W_t},
\]

\( \bar{W}_{jt} \) denotes the real wage rate for type \( j \) labor and \( \bar{W}_t \) denotes the real wage index

\[
\bar{W}_{jt} = \frac{W_{jt}}{P_t}, \quad \bar{W}_t = \frac{W_t}{P_t}.
\]

The price index \( P_t \) and the wage index \( W_t \) are defined later.

---

\(^9\)The price elasticity of demand may differ from the preference parameter \( \theta \) when households have imperfect information. The wage elasticity of labor demand may differ from the technology parameter \( \eta \) when firms have imperfect information.
3 Model setup – rational inattention

We now describe how decision-makers in firms and households allocate attention. A decision-maker who optimally allocates attention compares cost and benefit of paying attention. The benefit of paying attention to the current state of the economy is that actions get closer to the optimal actions 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 the decision-maker in a firm, we derive a simple expression for the loss in profit that a firm incurs if the firm’s actions 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’s actions deviate from the optimal actions under perfect information. We then state the attention problem of a household.

3.1 Loss in profit in the case of suboptimal actions

To derive the objective of a firm, one first has to state the demand function for an individual good and specify how a firm values profit in different states of the world.

We guess the following demand function for good $i$

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

Here $P_t = d(P_{1t}, \ldots, P_{Jt})$ is a price index and the function $d$ is homogenous of degree one, continuously differentiable, and symmetric. The coefficients $\vartheta$ and $\tilde{\theta}$ satisfy $\vartheta > 0$ and $\tilde{\theta} > 1$. We verify that the demand function has this form in equilibrium.

The economy described in the previous section is an incomplete markets economy with multiple owners of a firm. It is therefore in principle unclear how a firm values profit in different states of the world. For this reason, we assume a general stochastic discount factor: In period $-1$ the decision-maker in a firm values nominal profit in period $t$ using

$$Q_{-1,t} = \beta^t \Lambda(C_{1t}, \ldots, C_{Jt}) \frac{1}{P_t}, \quad (14)$$
where $C_{jt}$ is composite consumption by household $j$, $\Lambda$ is some twice continuously differentiable function and $P_t$ is the price index appearing in the demand function (13).

Substituting the production function (4)-(5) and the demand function (13) into the expression for profit (6) yields the profit function. Multiplying the profit function by the discount factor (14), summing over all periods, and taking the expectation conditional on information of the decision-maker in firm $i$ in period $-1$ yields the objective of the decision-maker in firm $i$ in period $-1$. Next, a log-quadratic approximation to this objective around the non-stochastic steady state yields a simple expression for the expected discounted sum of losses in profit from suboptimal actions:

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

where

$$x_t = \begin{pmatrix} p_{it} \\ \hat{l}_{i1t} \\ \vdots \\ \hat{l}_{i(J-1)t} \end{pmatrix}, \quad H = -\Lambda (C_1, \ldots, C_J) \tilde{P}_t C_i \begin{bmatrix} \tilde{\theta} \left(1 + \frac{1-\alpha}{\alpha} \tilde{\theta} \right) & 0 & \cdots & 0 \\ 0 & \frac{2\alpha}{\eta J} & \frac{\alpha}{\eta J} & \cdots & \frac{\alpha}{\eta J} \\ \vdots & \frac{\alpha}{\eta J} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{\alpha}{\eta J} \\ 0 & \frac{\alpha}{\eta J} & \cdots & \frac{\alpha}{\eta J} & \frac{2\alpha}{\eta J} \end{bmatrix},$$

and the optimal actions are given by

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

and

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

Here variables without time subscript denote values in the non-stochastic steady state and small variables denote log-deviations from the non-stochastic steady state.\(^\text{10}\)

Equations (17)-(18) are the usual log-linear optimality conditions familiar from the New Keynesian literature: The price that maximizes profit in period $t$ is a log-linear function of the price level, aggregate composite consumption, the real wage index, and total factor productivity. The relative

---

\(^{10}\) The non-stochastic steady state is characterized in online Appendix A. The derivation of expression (15) and equations (17)-(18) is in online Appendix B. The derivation only requires a weak regularity condition on the process \(\{x_t, x_t^*\}_{t=0}^{\infty}\).
input of type \( j \) labor that maximizes profit in period \( t \) depends only on the relative wage rate of type \( j \) labor. Furthermore, expression (15) implies that the profit loss in the case of a suboptimal action is only of second order.

3.2 Attention problem of the decision-maker in a firm

We now state the attention problem of the decision-maker in a firm. The decision-maker in firm \( i \) chooses in period \(-1\) the amount of attention allocated to the pricing decision and the labor mix decision as well as the law of motion for the price and the labor mix so as to maximize the expected discounted sum of profit net of the cost of paying attention:

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

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

\[
p^*_it = A_1(L) \varepsilon^A_i t + A_2(L) \varepsilon^R_i t + A_3(L) \varepsilon^I_i t,
\]

the law of motion for the actual actions

\[
p_{it} = B_1(L) \varepsilon^A_i t + C_1(L) \nu^A_i t + B_2(L) \varepsilon^R_i t + C_2(L) \nu^R_i t + B_3(L) \varepsilon^I_i t + C_3(L) \nu^I_i t,
\]

and the constraint on information flow

\[
\mathcal{I} \left( \left\{ p_{it}^s, p_{it}^R, p_{it}^I ; i_{i; (j-1)} t \right\}; \left\{ p_{it}^A, p_{it}^R, p_{it}^I, i_{i; (j-1)} t \right\} \right) \leq \kappa.
\]

Here \( A_s(L) \), \( B_s(L) \) and \( C_s(L) \) with \( s = 1, 2, 3 \) are infinite-order lag polynomials. Furthermore, the noise terms \( \nu^A_i t, \nu^R_i t, \nu^I_i t \) and \( \nu^{L_{ij}}_t \) follow Gaussian white noise processes with unit variance that are independent of fundamentals, independent of each other, and independent across firms.

The decision-maker in a firm who optimally allocates attention compares benefit and cost of paying attention. The benefit of paying more attention to the current state of the economy is that actions track more closely the optimal actions under perfect information. The cost of paying
attention can be thought of as time. The first term in objective (19) is the expected discounted sum of losses in profit when the law of motion for the actions differs from the law of motion for the optimal actions under perfect information, where the vector $x_t$ and the matrix $H$ are given by equation (16). The second term in objective (19) is the cost of attention. The parameter $\mu \geq 0$ is the per-period marginal cost of attention for the decision-maker in a firm and the variable $\kappa \geq 0$ is the amount of attention allocated to the pricing decision and the labor mix decision.

The optimal actions under perfect information are given by equations (20)-(21). Here we guess that the optimal price (17) is a linear function of current and past shocks in equilibrium. For example, $A_1(L)$ gives the response of the optimal price to aggregate technology shocks in equilibrium. We verify this guess. Furthermore, we express the optimal labor mix (18) as a function of relative wage rates.

The actual actions are given by equations (22)-(23). We allow for two types of deviations of the law of motion for the price from the law of motion for the optimal price under perfect information. First, the price may respond suboptimally (e.g., with dampening and delay) to fundamental shocks (i.e., $B_s(L) \neq A_s(L)$ for some $s$). Secondly, there may be noise in the agent’s price setting behavior (i.e., $C_s(L) \neq 0$ for some $s$). Of course, if there was no cost of absorbing and processing information, the agent would choose $B_s(L) = A_s(L)$ and $C_s(L) = 0$ for $s = 1, 2, 3$. The price set by the agent would then respond perfectly to all shocks in every period and there would be no noise in the agent’s pricing behavior. Similarly, the labor mix may respond suboptimally to shocks to relative wage rates (i.e., $\tilde{\eta} \neq \eta$) and there may be noise in the agent’s hiring behavior (i.e., $\chi > 0$).11

Finally, the left-hand side of the information flow constraint (24) measures the amount of information contained in the agent’s actions about the optimal actions under perfect information. It also measures the amount of information contained in the agent’s actions about the underlying shocks to the optimal actions. If the agent wants to take actions that track more closely the optimal actions under perfect information, the agent has to pay more attention.12

11 We put more structure on the labor mix decision than on the price setting decision by expressing the labor mix as a function of relative wage rates rather than fundamental shocks. We do this because from equation (23) we derive the labor demand function and a labor demand function specifies labor demand on and off the equilibrium path. By expressing the labor mix as a function of relative wage rates, we specify labor demand on and off the equilibrium path.

12 Following Sims (2003), the amount of information acquired and processed by the agent is quantified as uncertainty.
In the economy there are multiple shocks. An agent with unlimited attention would respond perfectly to all shocks. Our rational agent with costly attention chooses how much attention he devotes to aggregate technology, monetary policy, and local conditions. The more attention he devotes to aggregate technology shocks, the more closely $p^A_{it}$ tracks its perfect information target $p^{*A}_{it}$. The more attention he devotes to monetary policy shocks, the more closely $p^R_{it}$ tracks $p^{*R}_{it}$. The more attention he devotes to local conditions, the more closely $p^I_{it}$ tracks $p^{*I}_{it}$.

In Section 6 we formulate the attention problem of the decision-maker in a firm differently. We let the decision-maker choose the precision of signals subject to an information flow constraint. Afterwards, the decision-maker receives the signals and takes the best actions given those signals. When signals are about aggregate technology, monetary policy, and firm-specific productivity, we obtain identical results (i.e., the same law of motion for the price). Furthermore, when signals are about variables like the price level, total factor productivity, quantity sold, and the wage bill, we obtain similar results.

3.3 Loss in utility in the case of suboptimal actions

Let us turn to households. We first derive a simple expression for the utility loss that a household incurs if the household’s actions deviate from the optimal actions under perfect information. We then state the household’s attention problem.

We guess the following demand function for type $j$ labor

$$L_{jt} = \zeta \left( \frac{W_{jt}}{W_t} \right)^{-\tilde{\eta}} L_t. \quad (25)$$

Here $W_t = h(W_{1t}, \ldots, W_{Jt})$ is a wage index and the function $h$ is homogenous of degree one, continuously differentiable, and symmetric. The coefficients $\zeta$ and $\tilde{\eta}$ satisfy $\zeta > 0$ and $\tilde{\eta} > 1$. We verify that the labor demand function has this form in equilibrium.

Substituting the labor demand function (25), the flow budget constraint (3) and the consumption aggregator (2) into the period utility function (1) yields an expression for period utility that incorporates those three constraints. Multiplying by $\beta^t$, summing over all periods, and taking expectation conditional on information of household $j$ in period $-1$ yields the household’s objective.

A log-quadratic approximation to the household’s objective around the non-stochastic steady state reduction, where uncertainty is measured by entropy. The operator $\mathcal{I}$ is defined in online Appendix C.
yields the following expression for the expected loss in utility when the law of motion for the actions differs from the law of motion for the optimal actions under perfect information:

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

(26)

where

\[
x_t = \begin{pmatrix} \tilde{b}_{jt} \\ \tilde{w}_{jt} \\ \hat{c}_{1jt} \\ \vdots \\ \hat{c}_{I-1jt} \end{pmatrix}, \quad H_0 = -C_j^{1-\gamma} \begin{bmatrix} \gamma \omega_B^2 \left(1 + \frac{1}{\beta} \right) & \gamma \omega_B \tilde{\eta}_w \omega_W & 0 & \cdots & 0 \\ \gamma \omega_B \tilde{\eta}_w \omega_W & \tilde{\eta}_w \omega_W (\gamma \tilde{\eta}_w \omega_W + 1) & 0 & \cdots & 0 \\ 0 & 0 & \frac{2}{\beta} & \cdots & \frac{1}{\beta} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \frac{1}{\beta} & \frac{2}{\beta} \end{bmatrix},
\]  

(27)

and

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

(28)

Here \(\tilde{b}_{jt}\) denotes real bond holdings, \(\tilde{w}_{jt}\) denotes the real wage rate and \(\hat{c}_{ijt}\) denotes relative consumption of good \(i\). Furthermore, after the log-quadratic approximation to the household’s objective, the optimal actions under perfect information are given by

\[
\omega_B \hat{b}_{jt}^* = \frac{\omega_B}{\beta} \left( r_{t-1} - \pi_t + \hat{b}_{jt-1}^* \right) + \omega_W \frac{\tilde{\eta}}{\tilde{\eta} - 1} \left[ \left(1 - \tilde{\eta} \right) \tilde{w}_{jt}^* + \tilde{\eta} \hat{w}_t + l_t \right] + \omega_D \hat{d}_t - \omega T \hat{t}_t - c_{jt}^*,
\]  

(29)

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

(30)

\[
\tilde{w}_{jt}^* = \gamma c_{jt}^*,
\]  

(31)

and

\[
\hat{c}_{ijt}^* = -\theta \hat{p}_t.
\]  

(32)

Here \(E_t\) denotes the expectation operator conditioned on the entire history of the economy up to and including period \(t\) and the \(\omega\)’s denote steady state ratios given in online Appendix D.
Equations (29)-(32) are the usual log-linear optimality conditions familiar from the New Keynesian literature.\footnote{The derivation of expression (26) and equations (29)-(32) is in online Appendix D. The derivation only requires a weak regularity condition on the process $\{x_t, x_t^*\}_{t=0}^\infty$.}

Finally, the flow budget constraint (29) implies
\begin{equation}
\omega_B \left( \tilde{b}_{jt} - \tilde{b}^*_{jt} \right) = \frac{\omega_B}{\beta} \left( \tilde{b}_{jt-1} - \tilde{b}^*_{jt-1} \right) - \tilde{\eta}_W \left( \tilde{w}_{jt} - \tilde{w}^*_{jt} \right) - \left( c_{jt} - c^*_{jt} \right). \tag{33}
\end{equation}

Substituting this equation for the bond deviation into expression (26) and rearranging yields a simple expression for the expected utility loss in terms of consumption deviations, wage deviations and consumption mix deviations:
\begin{equation}
-C_j^{1-\gamma} \sum_{t=0}^\infty \beta^t \mathcal{E}_{j,-1} \begin{bmatrix}
\frac{\gamma}{2} \left( c_{jt} - c^*_{jt} \right)^2 + \tilde{\eta}_W \frac{W}{2} \left( \tilde{w}_{jt} - \tilde{w}^*_{jt} \right)^2 \\
\begin{bmatrix}
\hat{c}_{1jt} - \hat{c}^*_{1jt} \\
\vdots \\
\hat{c}_{I-1jt} - \hat{c}^*_{I-1jt}
\end{bmatrix}
\begin{bmatrix}
\frac{2}{\beta} \\
\vdots \\
\frac{2}{\beta}
\end{bmatrix}
\begin{bmatrix}
\hat{c}_{1jt} \\
\vdots \\
\hat{c}_{I-1jt}
\end{bmatrix}
\end{bmatrix}
\end{equation}

We will occasionally use this expression for the expected utility loss to provide intuition for results.\footnote{The derivation of expression (34) is also in online Appendix D and requires the same regularity condition as the derivation of expression (26).}

### 3.4 Attention problem of a household

We now state the attention problem of a household. Household $j$ chooses in period $-1$ the amount of attention allocated to the intertemporal consumption decision (how much to consume), the wage setting decision, and the consumption mix decision (which goods to consume) as well as the law of motion for composite consumption, the wage rate, and the consumption mix so as to maximize expected utility net of the cost of paying attention:
\begin{equation}
\max_{\kappa, B(L), C(L), \theta, \xi} \left\{ \sum_{t=0}^\infty \beta^t \mathcal{E}_{j,-1} \left[ \frac{1}{2} (x_t - x_t^*)' H_0 (x_t - x_t^*) + (x_t - x_t^*)' H_1 (x_{t+1} - x_{t+1}^*) \right] - \frac{\lambda}{1-\beta} \kappa \right\}, \tag{35}
\end{equation}

subject to equations (27)-(28), equation (33) for the bond deviation, the law of motion for the optimal actions under perfect information
\begin{equation}
c^*_jt = A_1 (L) \varepsilon^A_t + A_2 (L) \varepsilon^R_t, \tag{36}
\end{equation}
\[ \bar{w}_{jt}^* = \gamma c_{jt}^*, \quad (37) \]
\[ \hat{c}_{ijt}^* = -\theta \hat{p}_{it}, \quad (38) \]

the law of motion for the actual actions

\[ c_{jt} = B_1(L)\varepsilon_t^A + C_1(L)\nu_{jt}^A + B_2(L)\varepsilon_t^R + C_2(L)\nu_{jt}^R, \quad (39) \]
\[ \bar{w}_{jt} = \gamma c_{jt}, \quad (40) \]
\[ \hat{c}_{ijt} = -\hat{\theta}_\nu \hat{p}_{it} + \xi \nu_{ijt}, \quad (41) \]

and the constraint on information flow

\[ \mathcal{I} \left( \{c_{jt}^A, c_{jt}^R, c_{1jt}^*, \ldots, c_{I-1}^* \}; \{c_{jt}^A, c_{jt}^R, \hat{c}_{1jt}, \ldots, \hat{c}_{I-1} \} \right) \leq \kappa. \quad (42) \]

Here \( A_s(L), B_s(L) \) and \( C_s(L) \) with \( s = 1, 2 \) are finite-order lag polynomials. Furthermore, the noise terms \( \nu_{jt}^A, \nu_{jt}^R \) and \( \nu_{ijt}^I \) follow Gaussian white noise processes with unit variance that are independent of fundamentals, independent of each other, and independent across households.

A household who optimally allocates attention compares benefit and cost of paying attention. The benefit of paying more attention to the current state of the economy is that actions track more closely the optimal actions under perfect information. The first term in objective (35) is the expected utility loss when the law of motion for the actions differs from the law of motion for the optimal actions under perfect information. The cost of paying attention can be thought of as time. The parameter \( \lambda \geq 0 \) is the per-period marginal cost of attention for a household and the variable \( \kappa \geq 0 \) is the amount of attention allocated to the intertemporal consumption decision, the wage setting decision, and the consumption mix decision.

Equations (36), (38), (39), (41), and (42) mirror equations (20)-(24). For this reason, we do not comment on them again. The new features of the household’s problem compared to the firm’s problem are equation (33) and equations (37) and (40). We now comment on these new features. First, equation (33) gives the current bond deviation as a function of the past bond deviation, the current wage deviation, and the current consumption deviation. One has to compute the evolution of the bond deviation (independent of whether one uses expression (26) or (34) for stating the household’s objective) to verify that bond holdings satisfy the regularity condition that was used to derive expressions (26) and (34).
Second, the household makes two decisions that are closely related: the consumption decision and the wage setting decision. For simplicity, we assume that each household sets a real wage rate. Equation (37) gives the real wage rate that the household would set if the household had perfect information in every period. See equation (31). Next, equation (40) gives the actual real wage rate. Here we use the following result: It is always optimal for the household to equate the real wage rate to the marginal rate of substitution between consumption and leisure. In terms of log-deviations from the non-stochastic steady state, this intratemporal optimality condition reads $\tilde{w}_{jt} = \gamma c_{jt}$.\footnote{This intratemporal optimality condition can be easily derived from equation (33) and expression (34). First, any pair $(c_{jt}, \tilde{w}_{jt}) \in \mathbb{R}^2$ with the same $c_{jt} + \tilde{\eta}_W \tilde{w}_{jt}$ yields the same dynamics of bond holdings. See equation (33). Second, the unique pair $(c_{jt}, \tilde{w}_{jt}) \in \mathbb{R}^2$ that minimizes $(\gamma/2)(c_{jt} - c_{jt})^2 + (\tilde{\eta}_W/2)(\tilde{w}_{jt} - \tilde{w}_{jt})^2$ for a given $c_{jt} + \tilde{\eta}_W \tilde{w}_{jt}$ satisfies $\tilde{w}_{jt} = \gamma c_{jt}$.}

Furthermore, since the household chooses the real wage rate and consumption, the household can satisfy this intratemporal optimality condition independently of the nature of information flows to the household. Hence, a solution to the household’s attention problem has to satisfy equation (40).\footnote{Since composite consumption and the real wage rate contain the same information, it does not matter whether one uses the composite consumption behavior or the wage setting behavior in (42) to quantify the information flow to the household.} \footnote{We also solved the model assuming that households set nominal wage rates. See Section 6.}

### 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 optimal price, $p^*_t$, and the law of motion for optimal consumption, $c^*_t$. Second, we solve the firms’ attention problem (19)-(24) and the households’ attention problem (35)-(42). We turn each problem into a finite-dimensional problem by parameterizing each infinite-order lag polynomial $B_s(L)$ and $C_s(L)$ as a lag polynomial of a finite-order ARMA process. We then use a standard nonlinear optimization program to solve the firms’ problem and the households’ problem. Third, we aggregate across firms to obtain the price.
level, \( p_t = \frac{1}{T} \sum_{i=1}^{T} p_{it} \). We aggregate across households to obtain aggregate consumption and the real wage index, \( c_t = \frac{1}{J} \sum_{j=1}^{J} c_{jt} \) and \( \tilde{w}_t = \frac{1}{J} \sum_{j=1}^{J} \tilde{w}_{jt} \). Fourth, we compute the law of motion for the nominal interest rate from the monetary policy rule (which is log-linear) and \( y_t = c_t \). Finally, we compute the law of motion for the optimal price from equation (17) and the law of motion for optimal consumption from equation (30). If the law of motion for the optimal price or the law of motion for optimal consumption differs from our guess, we update the guess until a fixed point is reached.\(^{18}\)

### 4.2 Evaluation strategy

We pursue the following strategy to set parameter values and to evaluate the model. The strategy is in the spirit of impulse-response-matching 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 attention for the decision-maker in a firm, \( \mu \), and the marginal cost of attention for a household, \( \lambda \)) and all other parameters (non-rational-inattention parameters). We calibrate the non-rational-inattention parameters. We then solve the model for a grid of values of \( \mu \) and \( \lambda \). We select the pair of values of \( \mu \) and \( \lambda \) that minimizes the distance between the impulse responses to a monetary policy shock in the model and the impulse responses to a monetary policy shock in the VAR of Altig et al. (2011) (henceforth, ACEL VAR).

Afterwards, we evaluate the model by comparing: (i) the impulse responses to a monetary policy shock in the model and in the ACEL VAR, and (ii) the impulse responses to an aggregate technology shock in the model and in the ACEL VAR.\(^{19}\) A priori, it is unclear whether a DSGE model with rational inattention can match the empirical impulse responses to a monetary policy shock for any values of \( \mu \) and \( \lambda \). Furthermore, for the selected values of \( \mu \) and \( \lambda \), the model may or may not come close to matching the empirical impulse responses to an aggregate technology shock.

The DSGE literature has developed by matching impulse responses from VARs. We follow the same approach. We focus on the ACEL VAR, because from this VAR we get simultaneously empirical impulse responses to both aggregate shocks in our model. In the future, it will be important

---

\(^{18}\) See online Appendix E for more details regarding the solution method.

\(^{19}\) By “an aggregate technology shock” in Altig et al. (2011) we mean their neutral technology shock. In the VAR, Altig et al. (2011) make standard assumptions to identify a monetary policy shock and a neutral technology shock.
to estimate the model as in Smets and Wouters (2007), but we cannot undertake estimation before advances in the speed of computation allow us to solve the model faster than we currently do.

4.3 Calibration

The non-rational-inattention parameters can be calibrated in a straightforward way.

One period in the model is one quarter and therefore we assume $\beta = 0.99$. As is common in business cycle models, we set $\gamma = 1$ and $\alpha = 2/3$. To calibrate the parameters of the monetary policy rule we follow Taylor (1993) and let $\phi_{\pi} = 1.5$ and $\phi_{y} = 0.125$. We specify $\rho = 0.85$ because Clarida et al. (1999, p.1687) summarize the literature by writing that the estimates of this parameter “are typically on the order of 0.8 to 0.9.” We then compute a model-consistent measure of the monetary policy shock $\varepsilon_t^R$ from the data, $\varepsilon_t^R = r_t - \rho r_{t-1} - (1 - \rho) \left[ \phi_{\pi} \pi_t + \phi_{y} (y_t - y^n_t) \right]$. See equation (7).

We use the federal funds rate as a measure of the nominal interest rate, the difference of the log of the GDP deflator as a measure of inflation, and the difference between the log of real GDP and the log of real potential GDP reported by the Congressional Budget Office as a measure of the output gap. We employ quarterly data from 1982:1 to 2008:3. We choose this sample period because Altig et al. (2011) use this sample period. The standard deviation of the monetary policy shock obtained from this calculation is 0.0014. Therefore, we set the standard deviation of $\varepsilon_t^R$ in the model equal to 0.0014. To calibrate the parameters of the stochastic process for aggregate technology we use data on total factor productivity reported by Fernald (2014), from 1982:1 to 2008:3. 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 the autocorrelation of aggregate technology equal to 0.93 and the standard deviation of the aggregate technology shock $\varepsilon_t^A$ equal to 0.0064. It is worthwhile to point out that all these numbers are typical to the literature on business cycles.

To calibrate the parameters of the stochastic process for firm-specific productivity, we follow a common strategy in the menu cost literature (see, e.g., Golosov and Lucas (2007) and Nakamura and Steinsson (2008)): We choose the standard deviation of idiosyncratic shocks so that the size of price changes in the model matches the size of price changes in micro 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. Since our model is quarterly and $(0.67)^3 = 0.3$, we set the autocorrelation of firm-specific productivity equal to 0.3. Klenow and Kryvtsov (2008)
report that the median absolute size of price changes (excluding sale-related price changes) equals 9.7 percent in the micro data that the Bureau of Labor Statistics uses to compute the consumer price index. We therefore set the standard deviation of firm-specific productivity shocks so that the median absolute size of price changes equals 9.7 percent in our model. This choice yields a standard deviation of $\varepsilon_{it}$ equal to 0.23.

Two comments are in order. First, to compute the median absolute size of price changes in the model, we simulate prices when all decision-makers have perfect information and when decision-makers in firms and households face the selected values of $\mu$ and $\lambda$ (see Section 4.2). In either case, a standard deviation of $\varepsilon_{it}$ equal to 0.23 yields a median absolute size of price changes of 9.7 percent. The reason is simple. Price changes in the model are driven mostly by idiosyncratic shocks and firms therefore choose to track firm-specific productivity shocks almost perfectly under rational inattention. Second, the impulse responses of all variables in the model to both aggregate shocks are invariant to the calibration of the stochastic process for firm-specific productivity. This is because decision-makers face a constant marginal cost of attention. If the standard deviation of firm-specific productivity shocks falls, decision-makers will allocate less attention to firm-specific productivity shocks, but decision-makers will not change the amount of attention allocated to either aggregate shock, because neither the benefit nor the cost of paying attention to aggregate shocks has changed.

We choose the preference parameter $\theta$ and the technology parameter $\eta$ so as to match a price elasticity of demand of $\tilde{\theta} = 4$ and a wage elasticity of labor demand of $\tilde{\eta} = 4$. Choosing the preference parameter $\theta$ so as to match a particular price elasticity of demand is common practice in the literature on business cycle models with monopolistic competition because one cannot observe the preference parameter directly. When households have perfect information, calculating the value of the preference parameter that yields the targeted price elasticity of demand is easy because $\tilde{\theta} = \theta$. When households have limited attention, calculating the value of $\theta$ that yields the targeted $\tilde{\theta}$ is more complicated because $\tilde{\theta} < \theta$. We proceed as follows. The aggregate dynamics depend only on $\tilde{\theta}$ and $\tilde{\eta}$ (not on $\theta$ and $\eta$). Therefore, one can first solve for the aggregate dynamics of the model given the values of all other parameters and $\tilde{\theta} = \tilde{\eta} = 4$. One can then compute the value of $\theta$ that yields $\tilde{\theta} = 4$ and the value of $\eta$ that yields $\tilde{\eta} = 4$. We choose $\tilde{\theta} = 4$ because a price elasticity of demand of four is within the range of estimates of the price elasticity of demand in the Industrial
Organization literature. See Nakamura and Steinsson (2008, 2010). We choose $\tilde{\eta} = 4$ following Erceg et al. (2000). Furthermore, we set the number of consumption goods to $I = 100$ and the number of labor types to $J = 100$. The parameters $I$ and $J$ only affect the value of $\theta$ that yields $\tilde{\theta} = 4$ and the value of $\eta$ that yields $\tilde{\eta} = 4$.

We choose the ratio of wage income to consumption expenditure in the non-stochastic steady state, $\omega_W$, based on micro data from the Survey of Consumer Finances (SCF) 2007. We use micro data to calibrate this parameter, because we also want to investigate the effect of heterogeneity in this parameter on aggregate dynamics. From the SCF 2007, we compute median annual income, median annual wage income, and median net worth for households in the middle income quintile. (All variables are nominal.) From the first variable and the third variable we compute a proxy for annual consumption expenditure by applying the 2007 federal tax rate schedule “married filing jointly” to annual income and by deducting the number for annual savings that would leave real net worth constant at an annual inflation rate of 2.5 percent. Dividing annual wage income by this proxy for annual consumption expenditure yields a value of $\omega_W$ equal to 1.06. In Section 4.5, we also solve the model with five different types of households where the five values of $\omega_W$ are obtained by applying the same calculation to all five income quintiles. Furthermore, we solve the model with a single $\omega_W$ that is calculated from national accounts data. Finally, the parameters $\omega_D$, $\omega_T$, and $\omega_B$ do not affect the variables of interest. The parameters $\omega_D$ and $\omega_T$ do not appear in the households’ attention problem, and the parameter $\omega_B$ does not affect the solution to the households’ attention problem. The only role of $\omega_B$ is to turn consumption and wage deviations into bond deviations in equation (33) and to translate those bond deviations back into consumption deviations in objective (35).

Holding the non-rational-inattention parameters constant at the selected values, we solve for the aggregate dynamics for a grid of values of the rational-inattention parameters $\mu$ and $\lambda$. We find that $\mu = 0.0006 \times \Lambda(C_1, \ldots, C_J) \hat{P}_i C_i$ and $\lambda = 0.0006 \times C_{ij}^{1-\gamma}$ minimizes the sum of squared differences between the impulse responses of aggregate 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 ACEL VAR.\textsuperscript{20}

\textsuperscript{20}The fact that $\mu = 0.0006 \times \Lambda(C_1, \ldots, C_J) \hat{P}_i C_i$ and $\lambda = 0.0006 \times C_{ij}^{1-\gamma}$ simply means that we obtain a marginal cost of attention for a firm equal to 0.0006 after we divide the firms’ objective (19) by the term $\Lambda(C_1, \ldots, C_J) \hat{P}_i C_i$ that appears in the equation for $H$ and that we obtain a marginal cost of attention for a household equal to 0.0006.
We refer to the model with these parameter values as the baseline economy.

4.4 Baseline economy: results

The main result is that the model matches the empirical impulse responses to a monetary policy shock for the selected values of $\mu$ and $\lambda$. Furthermore, for the same values of $\mu$ and $\lambda$, the model matches the empirical impulse responses to an aggregate technology shock.

Let us begin with monetary policy shocks. The left column of Figure 1 depicts the impulse responses of output and inflation to a monetary policy shock in the baseline economy and in the ACEL VAR. For comparison, Figure 1 also shows the analogous impulse responses in the DSGE model of Altig et al. (2011) (henceforth, ACEL DSGE).\footnote{The ACEL DSGE is a standard medium-sized New Keynesian DSGE model, built to mimic the impulse responses from the ACEL VAR. Altig et al. (2011) assume in their VAR and in their model that in any quarter pricing and consumption decisions are made before this quarter’s monetary policy shock is realized. To facilitate comparison, we make the same timing assumption in the baseline economy.}

Table 1 reports the second moments conditional on a monetary policy shock in our model, in the ACEL VAR, and in the ACEL DSGE.

The impulse response of output to a monetary policy shock in the baseline economy matches very well the analogous impulse response in the ACEL VAR: the size and shape of the response in the model are very close to the size and shape of the empirical response. In particular, the response in the baseline economy is hump-shaped, like in the data, and the autocorrelation of output growth conditional on a monetary policy shock in the baseline economy is about 0.5, like in the data. Note also that the response in the baseline economy is essentially the same as the response in the ACEL DSGE.

The impulse response of inflation to a monetary policy shock in the baseline economy matches fairly well the analogous impulse response in the ACEL VAR: the size of the response in the baseline economy is correct, and the response is as persistent as in the data (the autocorrelation of inflation conditional on a monetary policy shock is about 0.9 both in the baseline economy and in the ACEL VAR). However, the response in the baseline economy is monotonic whereas the response in the ACEL VAR is hump-shaped. Two comments are in order. First, although the response in the baseline economy is not hump-shaped, it lies within the confidence interval from the ACEL VAR. Second, the reason why the ACEL DSGE and some other New Keynesian models can produce a
hump-shaped response of inflation to a monetary policy shock is that these models assume an extra source of slow adjustment, price indexation.

For the model to match the empirical impulse responses to a monetary policy shock, it is essential that the marginal cost of attention is positive for firms and households. This point is clear from Table 1 which compares the second moments conditional on a monetary policy shock in the baseline economy with the analogous objects in the economy with perfect information ($\mu = \lambda = 0$) and in the economy in which only decision-makers in firms are subject to rational inattention ($\mu = 0.0006*\Lambda (C_1, \ldots, C_J) \hat{P}_iC_i$ and $\lambda = 0$). In the economy with perfect information, a monetary policy shock has no real effects. If only decision-makers in firms are subject to rational inattention, the impulse response of output to a monetary policy shock is monotonic rather than hump-shaped, and the autocorrelation of output growth conditional on a monetary policy shock is negative rather than positive.

The right column of Figure 1 and Table 2 show the results in the case of aggregate technology shocks. The impulse response of output to an aggregate technology shock in the baseline economy is close to the analogous impulse response in the ACEL VAR. One difference is that output growth conditional on an aggregate technology shock is autocorrelated in the baseline economy, whereas output is approximately a random walk at the point estimate in the ACEL VAR. However, the confidence interval from the ACEL VAR is consistent with autocorrelation in output growth. Furthermore, output growth conditional on an aggregate technology shock is autocorrelated also in the ACEL DSGE.

The impulse response of inflation to an aggregate technology shock in the baseline economy is also close to the analogous impulse response in the ACEL VAR. Inflation responds strongly on impact to an aggregate technology shock in the data, and the model fits this observation. By contrast, the ACEL DSGE fails to fit this observation. In the ACEL DSGE, the response of inflation to an aggregate technology shock is essentially the same as the response of inflation to a monetary policy shock.22

We conclude that the rational-inattention DSGE model matches the empirical impulse responses to a monetary policy shock for the selected values of $\mu$ and $\lambda$. Furthermore, for the same values of

---

22 In a detailed VAR study, Paciello (2011) documents that it is a robust finding that inflation adjusts much faster to aggregate technology shocks than to monetary policy shocks.
μ and λ, the model matches the empirical impulse responses to an aggregate technology shock.23

So far we discussed conditional second moments. The ACEL VAR implies a standard deviation of output growth due to monetary policy shocks of 0.0012 (Table 1) and a standard deviation of output growth due to aggregate technology shocks of 0.0025 (Table 2). Together these numbers imply a standard deviation of output growth of 0.0028. By contrast, the unconditional standard deviation of output growth in the data is 0.0065. This means that the variation in output in the data is generated by more than these two shocks, and the same is true about the variation in inflation in the data.24 The upshot is that any DSGE model that matches the moments conditional on monetary policy shocks and on aggregate technology shocks needs more than these two shocks to match the unconditional moments. In the future, it will be worthwhile to add more shocks to our model to fit simultaneously the conditional moments and the unconditional moments. Adding more shocks to the model will not change the answers to the following questions. First, for which values of the marginal costs of attention does the model match the empirical impulse responses to a monetary policy shock? Second, for these values of the marginal costs of attention, does the model also match the empirical impulse responses to an aggregate technology shock? The reason why the answers to these questions will not change is that adding more shocks does not change the marginal benefit of paying attention to the shocks already present in the model. Similarly, moving from a linear to a non-linear attention cost function will also not change the answers to these questions.

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 households and decision-makers in firms.

4.5 Understanding the behavior of households

To develop intuition for the households’ behavior, it is useful to calculate the utility loss associated with a simple deviation from the optimal consumption path. Cochrane (1989) gives a similar example to illustrate that deviations from the optimal consumption path carry small utility losses.

23 In a previous version of the paper, we selected the values of μ and λ to match the impulse responses to a monetary policy shock in the estimated DSGE model in Smets and Wouters (2007). The conclusions were the same as here.

24 The unconditional standard deviation of inflation in the data is 0.0026. When computing the unconditional second moments in the data, we use the difference of the log of real GDP per capita as a measure of output growth and the difference of the log of the GDP deflator as a measure of inflation. The data run from 1982:1 to 2008:3. Altig et al. (2011) use the same data.
Suppose a contractionary monetary policy shock hits the economy in period zero, a household reduces consumption by 0.1 percent, and the optimal consumption reduction is 0.6 percent, i.e., the deviation from the optimal consumption path on impact of the shock equals half a percent \((c_{j0} - c^*_j = 0.005)\). In the next quarter, the household offsets the effect of this consumption deviation on the present-value budget constraint by consuming the required amount less than dictated by the optimal consumption plan \((c_{j1} - c^*_j = -1.01 \times 0.005\) if one assumes a real interest rate of one percent). The utility loss associated with this double deviation from the optimal consumption path can be easily calculated from equations (26)-(27), because \(\tilde{b}_{jt} - \tilde{b}^*_j = -\omega B \times 0.005\) in period zero and \(\tilde{b}_{jt} - \tilde{b}^*_j = 0\) in all following periods. The utility loss equals \(\frac{1}{2} C^{1-\gamma} (1 + 1.01) (0.005)^2\). Dividing by the marginal utility of consumption \(C_j\) expresses the utility loss in consumption equivalent. Dividing further by \(C_j\) expresses it as a fraction of steady-state consumption. Hence, this simple double deviation from the optimal consumption path carries a utility loss of about \((\gamma/40,000)\) of steady-state consumption.

Furthermore, this is the size of a deviation from the optimal consumption path that is sufficient to match the data. Figure 2 shows the actual consumption response and the optimal consumption response to a one-standard-deviation monetary policy shock at the fixed point of the baseline economy. The initial consumption deviation equals 0.5 percent. The only two differences between the deviation from the optimal consumption path shown in Figure 2 and the deviation from the optimal consumption path discussed in the previous paragraph are: the household consumes too much for seven quarters instead of one quarter (which increases the utility loss), and the household spreads out the adjustment that restores the budget constraint over many quarters rather than concentrating it in one quarter (which reduces the utility loss). The second effect dominates slightly. The consumption pattern shown in Figure 2 carries a utility loss of \((\gamma/40,954)\) of steady-state consumption. Since this utility loss is so small, a household is willing to tolerate this consumption deviation (i.e., chooses not to reduce the deviation by paying more attention) even if the marginal cost of attention is small.

A household also incurs utility losses due to suboptimal labor market decisions. In any period, the combined utility loss equals \(\frac{1}{2} \gamma\) times the squared consumption deviation plus \(\frac{1}{2} \tilde{\eta} \omega W\) times the squared wage deviation (see objective (34)). The wage deviation equals \(\gamma\) times the consumption deviation. Thus, the two ratios for the utility loss given in the previous two paragraphs have
\((\gamma + \tilde{\eta}\omega_W\gamma^2)\) instead of \(\gamma\) in the numerator once we take into account that households also incur utility losses due to suboptimal labor market decisions. In addition, consumption and wage deviations are generated by noise in the perception of monetary policy shocks. We arrive at the following total utility loss due to suboptimal tracking of monetary policy shocks: For the impulse responses of consumption and wages to monetary policy shocks and to noise in the baseline economy (see Figure 2) and for the baseline parameter values \((\gamma = 1, \tilde{\eta} = 4, \omega_W = 1.06)\), the expected per-period utility loss due to suboptimal tracking of monetary policy shocks equals \((1/5,907)\) of steady-state consumption. Since this expected utility loss is so small, households find it optimal to pay little attention to monetary policy shocks even if the marginal cost of attention is small.\(^{25}\)

Consumption patterns with the same absolute deviations from the optimal consumption path carry the same utility loss (see objective (34)). For example, an under-response of consumption on impact of a shock carries the same utility loss as an over-response of consumption on impact of a shock. The reason why the model generates under-responses of actions to shocks is that attention is costly and inattention generates delayed responses. What do households pay close attention to? In the baseline economy households pay seven times more attention to the relative price of an individual good than to monetary policy. Relative prices of goods are highly volatile, like in the data, and households thus find it important to be aware of those fluctuations. Of course, in the real world households need to track many more things than in a model. What matters for any given attention choice is the marginal benefit of paying attention and the opportunity cost of paying attention. The opportunity cost of paying attention is formalized by the marginal cost of attention, \(\lambda\). This marginal cost of attention can be interpreted as the Lagrange multiplier on a household’s attention constraint in an environment where households face all the decisions and shocks that households actually face in reality.

Next, we investigate how a household’s optimal attention allocation depends on \(\gamma, \tilde{\eta}, \omega_W\), and the utility function. First, we calibrate the ratio of wage income to consumption expenditure in the non-stochastic steady state, \(\omega_W\), using Bureau of Economic Analysis national accounts data (instead of micro data). The time average of the ratio of compensation of employees to personal

\(^{25}\)Figure 2 also shows the impulse responses of consumption to aggregate technology shocks and to noise in the perception of aggregate technology shocks in the baseline economy. The expected per-period utility loss due to suboptimal tracking of aggregate technology shocks in the baseline economy equals \((1/7,356)\) of steady-state consumption.
consumption expenditures in the period 1982-2008 (the sample of Altig et al. (2011)) equals 0.85. In the economy with $\omega_W = 0.85$ (instead of $\omega_W = 1.06$), households have even less incentive to pay attention to aggregate conditions, but multiplying the households’ marginal cost of attention by the ratio $[(\gamma + \tilde{\eta}\omega_W^{new}_W \gamma^2) / (\gamma + \tilde{\eta}\omega_W^{old}_W \gamma^2)]$ yields identical impulse responses as in the baseline economy. The same argument applies to a reduction in $\tilde{\eta}$. Hence, changes in $\omega_W$ or $\tilde{\eta}$ do not affect the model’s ability to match the empirical impulse responses of output and inflation to shocks. They only affect the value of $\lambda$ at which the best match is attained. Second, we investigate the effect of heterogeneity in the parameter $\omega_W$ across households on aggregate dynamics. We solve the model with five different types of households, where the five values of $\omega_W$ are obtained by applying the same calculation as before to all five income quintiles in the SCF 2007 (see Section 4.3). In this model with ex-ante heterogeneity, households with a higher $\omega_W$ pay more attention to aggregate conditions, but the impulse responses of aggregate consumption are essentially the same as in the baseline economy, where the parameter $\omega_W$ is calibrated using data for the middle income quintile. See online Appendix F. Third, increasing the coefficient of relative risk aversion always reduces a household’s incentive to pay attention to the macroeconomy so long as $C_j \geq 1$. This can be seen from equations (30)-(31) and expression (34). Fix for a moment the allocation of attention of households and consider an increase in $\gamma$ by a factor of two. The responses of optimal composite consumption to shocks are scaled down by a factor of two. To see this, solve equation (30) forward. For a given allocation of attention, the responses of actual composite consumption to shocks are also scaled down by a factor of two. Thus, the term $\left(c_{jt} - c_{jt}^*\right)^2$ in expression (34) is scaled down by a factor of four and the term $\tilde{\gamma}/2 \left(c_{jt} - c_{jt}^*\right)^2$ is scaled down by a factor of two. In other words, consumption deviations become smaller and more costly, and the first effect dominates. Furthermore, increasing $\gamma$ does not change the responses of the optimal real wage rate to shocks. To see this, substitute equation (31) into equation (30). Thus, for a given allocation of attention, the responses of the actual real wage rate to shocks do not change and the term $\tilde{\omega}_W/2 \left(\tilde{w}_{jt} - \tilde{w}_{jt}^*\right)^2$ in expression (34) does not change. Finally, the constant $C_j^{1-\gamma}$ in expression (34) is non-increasing in $\gamma$ so long as $C_j \geq 1$. Hence, increasing the coefficient of relative risk aversion always reduces a household’s incentive to pay attention to the macroeconomy so long as $C_j \geq 1$.

Fourth, we solved the model with external habit formation in consumption (i.e., the period utility function is $\ln(C_{jt} - \varphi_T C_{t-1}) - \varphi L_{jt}$, where $\varphi$ is the habit parameter and $(1/J)C_{t-1}$ is
average composite consumption in the previous period). In the baseline economy with $\rho = 0$, the ratio of actual consumption to optimal consumption on impact of a monetary policy shock equals 12 percent. See Figure 2. With $\rho = 0.3$ and $\rho = 0.6$, this ratio increases somewhat to 21 percent and 30 percent, respectively. The impulse response of aggregate output to a monetary policy shock shows even more dampening and delay than in the baseline economy, because now inattention and habit formation cause slow adjustment.²⁶

### 4.6 Understanding the behavior of firms

In the baseline economy, most of the variation in the profit-maximizing price (the price that the decision-maker would set if he had perfect information) is due to firm-specific productivity shocks, some of its variation is due to aggregate technology shocks, and little of its variation is due to monetary policy shocks. The decision-maker who optimally allocates attention therefore allocates a lot of attention to firm-specific productivity, some attention to aggregate technology, and little attention to monetary policy. Specifically, of the total attention devoted to the price-setting decision, the decision-maker allocates 93 percent to firm-specific productivity, 5 percent to aggregate technology, and 2 percent to monetary policy. Prices in the baseline economy thus respond very quickly to firm-specific productivity shocks, fairly quickly to aggregate technology shocks, and slowly to monetary policy shocks. Hence, the model matches the empirical finding by Boivin et al. (2009) and Maćkowiak et al. (2009) that prices respond very quickly to disaggregate shocks, and the model produces a much stronger response of inflation to an aggregate technology shock than to a monetary policy shock, consistent with the VAR evidence discussed in Section 4.4.²⁷

The endogenous attention allocation also implies that monetary policy has strong real effects while profit losses are small. In any model with a price setting friction, firms experience profit

²⁶There are two effects on households’ incentive to pay attention to the macroeconomy when the habit parameter increases. Optimal consumption responds more slowly to shocks (which reduces the incentive to pay attention), and a given deviation between actual consumption and optimal consumption becomes more costly (which increases the incentive to pay attention). The second effect slightly dominates.

²⁷In the baseline economy, prices respond almost perfectly to firm-specific productivity shocks. The response of the actual price on impact of a firm-specific productivity shock equals 99% of the response of the profit-maximizing price on impact of the same shock. When one reduces the standard deviation of firm-specific productivity shocks by a factor of 2 and 3, this number still equals 98% and 95%, respectively. In all three cases, the actual response equals 100% of the profit-maximizing response one period after the shock.
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 \(1/1,700\) of a firm’s steady-state revenue.\(^{28}\) The analogous number in the Calvo-with-habit model is \textit{ninety times larger}. The Calvo-with-habit model is a popular small DSGE model used to study the effects of monetary policy. The model has the same structure as the model in this paper except that decision-makers in firms face the Calvo friction in price setting, households are subject to external habit formation in consumption, and all agents have perfect information. The reason why profit losses are so much smaller in the rational inattention model than in the Calvo-with-habit model is that in the rational inattention model prices respond slowly only to unimportant shocks but quickly to important shocks.\(^{29}\)

The variation in the profit-maximizing price generated by the different shocks in equilibrium is determined by the non-rational-inattention parameters and feedback effects (the profit-maximizing price of a firm depends on the actions of other agents in the economy). The non-rational-inattention parameters have been chosen to match key features of the data: the large average absolute size of price changes in micro data and the small standard deviation of the innovation in the monetary policy rule. We now turn to the feedback effects.

### 4.7 Feedback effects

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 all households have perfect information. The profit-maximizing price response of this single firm to a monetary policy shock is several times larger in absolute terms than in the baseline economy. As a result, this single firm chooses to allocate five times more attention to monetary policy than in the baseline economy.

Next, suppose that all other firms also become subject to rational inattention. The profit-

\(^{28}\)Of this profit loss, about 5% is due to imperfect tracking of monetary policy, 20% is due to imperfect tracking of aggregate technology, and 75% is due to imperfect tracking of firm-specific productivity.

\(^{29}\)We compute profit losses in the Calvo-with-habit model having assumed the same values for preference, technology, and monetary policy parameters as in the baseline economy and having chosen values for the Calvo parameter and the habit parameter in the same way as values of the rational inattention parameters in the baseline economy. Our procedure yields a Calvo parameter of 0.9 and a habit parameter of 0.85.
maximizing price is given by
\[ p_{it} = p_t + \frac{1-\alpha}{\alpha} c_t^\gamma - \frac{1}{1 + \frac{1-\alpha}{\alpha} \phi} a_t - \frac{1}{1 + \frac{1-\alpha}{\alpha} \phi} a_{it}. \] (43)

When all firms absorb information slowly, the price level falls less after a contractionary monetary policy shock, but also consumption falls after a contractionary monetary policy shock (monetary policy has real effects). This changes the response of the profit-maximizing price to a monetary policy shock. With our parameterization, the profit-maximizing price moves much less after a monetary policy shock and hence firms choose to pay much less attention to monetary policy.

Finally, suppose that all households also become subject to rational inattention (the baseline economy). Additional feedback effects arise. The impulse response of consumption to a monetary policy shock becomes hump-shaped instead of monotonic (see Section 4.5) and thus the impulse response of the profit-maximizing price 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 again to much less attention to monetary policy.

There are also feedback effects between households. Suppose that all firms and a single household are subject to rational inattention, while all other households have perfect information. The optimal consumption response to a monetary policy shock is smaller in absolute terms than in the baseline economy, and therefore the single household with a positive marginal cost of attention pays less attention to monetary policy than in the baseline economy. The reason is the monetary policy rule. When a contractionary monetary policy shock arrives, the nominal interest rate increases. The size of this increase is always attenuated in equilibrium because output decreases and output enters the monetary policy rule. This attenuation effect is stronger when all other households have perfect information.

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.\textsuperscript{30} 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 \( c_t \), while the profit-maximizing price response to an aggregate technology shock depends in addition on an exogenous variable, \( a_t \). See equation (43). The effect of \( a_t \) on the profit-maximizing price is present with constant strength no matter what attention choices agents make.\textsuperscript{31} 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 (43) 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.

5 Great Moderation

In this section we describe an important property of the model. When we use the model to conduct an experiment (i.e., we change a parameter value), the outcome of the experiment is very different from the outcome of the same experiment in a New Keynesian model and in a model with exogenous dispersed information. The reason is intuitive. When a parameter value changes, decision-makers in this model reallocate attention. The reallocation of attention turns out to have a critical effect on the outcomes of experiments.

In an earlier version of the paper, we illustrated this property of the model with several experiments (see Maćkowiak and Wiederholt (2011)). Here we focus on a single data-driven experiment. We aim to show that a rational inattention DSGE model can help us understand the effects of structural change in the real world. Probably the most natural episode of structural change to study with this model is “the Great Moderation,” i.e., the well-known fact that the volatility of

\textsuperscript{30} Analogously, the optimal attention allocation of a household depends on the attention allocation of other households and on the attention allocation of firms.

\textsuperscript{31} 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, \( a_{it} \).
many macro variables declined in the 1980s and remained low. In the period 1960:1-1981:4, the
standard deviation of output growth was 0.0107 and the standard deviation of inflation was 0.0072.
In the period 1982:1-2008:3 (the sample of Altig et al. (2011)), the standard deviation of output
growth declined from 0.0107 to 0.0065 and the standard deviation of inflation declined from 0.0072
to 0.0026.32

A popular hypothesis is that the Great Moderation was caused by a change in policy. We model
the shift to anti-inflationary monetary policy by the Federal Reserve under the chairmanship of
Paul Volcker as an increase in the coefficient on inflation in the monetary policy rule. When
we raise $\phi_\pi$ from 1.01 to 1.5 (the baseline economy), the variance of output growth conditional
on aggregate technology shocks decreases and the variance of inflation conditional on aggregate
technology shocks decreases. Interestingly, in a basic New Keynesian model, this is not the case:
the variance of output growth conditional on aggregate technology shocks increases.

The intuition is the following. Fix for a moment the allocation of attention of decision-makers
in firms and households. Furthermore, focus on aggregate technology shocks. When monetary
policy reacts more aggressively to inflation, the nominal interest rate mimics more closely the real
interest rate that implements the efficient output level in a world where households have perfect
information. This effect decreases deviations of output from efficient output, i.e., the volatility of
the output gap falls and the volatility of output rises. Furthermore, the volatility of inflation falls.
However, in the rational inattention DSGE model there is another effect. When monetary policy
reacts more aggressively to inflation, the price level becomes more stable implying that decision-
makers in firms decide to pay less attention to the macroeconomy. This effect increases deviations
of output from efficient output, i.e., the volatility of the output gap rises and the volatility of output
falls. When we raise $\phi_\pi$ from 1.01 to 1.5, the second effect dominates and thus the volatility of
output falls.

The second effect is absent from a New Keynesian model and from a model with exogenous
dispersed information. Hence, in a basic New Keynesian model raising $\phi_\pi$ from 1.01 to 1.5 increases
the volatility of output due to aggregate technology shocks. For this reason, change-in-policy
explanations of the Great Moderation in the New Keynesian literature typically involve moving

\footnote{As in Section 4, we use the difference of the log of real GDP per capita as a measure of output growth and the
difference of the log of the GDP deflator as a measure of inflation.}
from an indeterminacy region to a determinacy region of the parameter space or involve other shocks. See, for example, Lubik and Schorfheide (2004).

Survey data on expectations provide direct evidence of the reallocation of attention predicted by the model. Coibion and Gorodnichenko (2012) study survey data on expectations finding that the degree of attention to the aggregate economy rose during the turbulent 1970s and fell during the subsequent calm period, consistent with our model.

6 Extensions

In this section we summarize the results from several extensions of the model. The details of the extensions are in online Appendix G.

To begin, we vary assumptions concerning information flows. So far we assumed that the decision-maker in a firm chooses a law of motion for the actions, subject to a constraint on the information content of the actions. We now state the attention problem of the decision-maker in a firm using signals. We assume that the decision-maker 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 receives the signals and takes the optimal actions given the signal realizations.

We solve this problem for an individual firm assuming that aggregate variables and relative wage rates are as in the equilibrium of the baseline economy. We compare the solution for the firm’s price and labor mix with the solution of the problem (19)-(24). When the signals are of the form “an optimal action under perfect information plus i.i.d. noise,” the two solutions are very similar. When we allow for the noise in the signals to be serially correlated, the two solutions are identical: the firm’s price and labor mix are the same as in the baseline economy. We then relax the assumption that paying attention to aggregate technology, paying attention to monetary policy, and paying attention to firm-specific productivity are independent activities. For example, we consider the case in which the decision-maker in a firm observes signals concerning the price level, total factor productivity, last period’s quantity sold, last period’s wage bill, and the relative wage rates. These variables are driven by multiple shocks, and thus it is 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 the solution is similar to the solution of the problem (19)-(24). The reason is that the decision-maker decides to pay close attention to those variables that are mainly driven by firm-specific productivity shocks and aggregate technology shocks.33

As another extension, we solve the model assuming that households set nominal wage rates instead of real wage rates. The main change is that rational inattention by households now also causes deviations from the households' intratemporal optimality condition stating that the real wage rate should equal the marginal rate of substitution between consumption and leisure. This creates small changes in the impulse responses of output and inflation to a monetary policy shock at given values of the marginal costs of attention. See online Appendix G. We chose to present the results with households setting real wage rates here, because this version of the model exhibits in the most transparent way the effects of rational inattention on the consumption-saving decision.

7 Conclusions and further research

Making good decisions in an environment with shocks requires attention. Attention is a scarce resource and decision-makers choose how to allocate it. We solve a DSGE model with rational inattention. Households and decision-makers in firms have limited attention and optimally allocate attention. Rational inattention is the only source of slow adjustment. The model matches the empirical impulse responses to monetary policy shocks and aggregate technology shocks. At the same time, profit losses and utility losses from inattention are very small. We conclude that the slow adjustment in macroeconomic data that is usually modelled with various forms of adjustment costs may actually have a different source: rational inattention.

There are several directions for future research. To not complicate the households’ problem further, we did not allow for trade in equity among households. It will be interesting to introduce equity markets and to study the asset pricing implications of the model. It is well known that standard DSGE models need high risk aversion to match jointly the high equity premium and the low contemporaneous covariance between consumption growth and equity returns in the data. Macroeconomists have responded to this fact by changing preferences in DSGE models (see, e.g.,

33 We include last period’s quantity sold and wage bill in the signal vector because we do not know how the firm can observe current period’s quantity sold and wage bill before setting the price. We studied a large number of variations of this signal vector and obtained similar results.
Tallarini (2000) and Boldrin et al. (2001)) and by introducing limited stock market participation and heterogeneity in preferences (see Guvenen (2009)). Slow adjustment of consumption to shocks due to inattention has been proposed as another possible explanation of the equity-premium puzzle. When consumption responds with delay to shocks, the riskiness of equity no longer shows up in a high contemporaneous covariance between consumption growth and equity returns (see, e.g., Gabaix and Laibson (2002)). It will be interesting to study to what extent the slow consumption responses in the rational inattention DSGE model can explain the equity-premium puzzle. Furthermore, it will be interesting to study the effect of limited stock market participation or recursive utility on the incentives to pay attention. We conjecture that the marginal cost of attention for households will have to be higher to generate impulse responses of aggregate consumption to shocks that are similar to the impulse responses in the baseline economy.34

It will also be interesting to introduce capital in the model. It seems reasonable to assume that firms own the capital stock and decision-makers in firms make the investment decision. In this case, the attention problem of households does not change, but the attention problem of firms changes. It will be interesting to study whether there exists a value for the marginal cost of attention for decision-makers in firms at which a rational inattention DSGE model matches simultaneously empirical responses of prices to shocks and empirical responses of investment to shocks.

Finally, it will be interesting to study optimal monetary policy and optimal fiscal policy in rational inattention DSGE models. Paciello and Wiederholt (2014) study optimal monetary policy in a DSGE model with rational inattention by decision-makers in firms. It is an open question how rational inattention by households affects optimal policy.

34 Luo and Young (2014) study the interaction of recursive utility and inattention in a pure exchange economy with two assets (a risky asset and a riskless asset). They find that even a small preference for early resolution of uncertainty reduces significantly the demand for the risky asset when agents have limited information-processing capacity. However, they do not study the effect of recursive utility on the incentives to pay attention.
References


[34] Paciello, Luigi (2011): “Does Inflation Adjust Faster to Aggregate Technology Shocks than to Monetary Policy Shocks?” Journal of Money, Credit and Banking, 43(8), 1663-1684.


Figure 1: Impulse responses in the baseline economy and in Altig et al. (2011)

Sources: Altig et al. (2011) and own calculations.
An impulse response equal to 1 means a 1 percent deviation from the non-stochastic steady state.
The ACEL VAR impulse responses are shown with 95 percent confidence intervals.
**Table 1: Second moments conditional on a monetary policy shock**

<table>
<thead>
<tr>
<th></th>
<th>Output growth</th>
<th></th>
<th>Inflation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard deviation</td>
<td>Autocorrelation</td>
<td>Standard deviation</td>
<td>Autocorrelation</td>
</tr>
<tr>
<td><strong>Data (ACEL VAR)</strong></td>
<td>0.0012</td>
<td>0.52</td>
<td>0.0003*</td>
<td>0.93*</td>
</tr>
<tr>
<td><strong>Baseline economy</strong></td>
<td>0.0009</td>
<td>0.54</td>
<td>0.0003</td>
<td>0.92</td>
</tr>
<tr>
<td><strong>ACEL DSGE</strong></td>
<td>0.0009</td>
<td>0.51</td>
<td>0.0003</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>Memo items:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Perfect information</strong></td>
<td>0</td>
<td>N/A</td>
<td>0.0053</td>
<td>0</td>
</tr>
<tr>
<td><strong>Only firms subject to RI</strong></td>
<td>0.0050</td>
<td>-0.15</td>
<td>0.0007</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Sources: Altig et al. (2011) and own calculations.

* The impulse response of inflation to a monetary policy shock in the ACEL VAR is choppy. To remove the effects of the choppiness on the second moments, we report the standard deviation and autocorrelation of the four-quarter-moving-average of this impulse response.

** Baseline economy is this paper's DSGE model with rational inattention with the parameter values given in Section 4.3.

*** Perfect information is the same as baseline economy except that μ=λ=0.

**** Only firms subject to RI, where RI stands for rational inattention, is the same as baseline economy except that λ=0.

---

**Table 2: Second moments conditional on an aggregate technology shock**

<table>
<thead>
<tr>
<th></th>
<th>Output growth</th>
<th></th>
<th>Inflation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard deviation</td>
<td>Autocorrelation</td>
<td>Standard deviation</td>
<td>Autocorrelation</td>
</tr>
<tr>
<td><strong>Data (ACEL VAR)</strong></td>
<td>0.0025</td>
<td>0.07</td>
<td>0.0009</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Baseline economy</strong></td>
<td>0.0021</td>
<td>0.66</td>
<td>0.0011</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>ACEL DSGE</strong></td>
<td>0.0016</td>
<td>0.75</td>
<td>0.0002</td>
<td>0.96</td>
</tr>
<tr>
<td><strong>Memo items:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Perfect information</strong></td>
<td>0.0065</td>
<td>-0.04</td>
<td>0.0032</td>
<td>0.12</td>
</tr>
<tr>
<td><strong>Only firms subject to RI</strong></td>
<td>0.0038</td>
<td>0.07</td>
<td>0.0007</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Sources: Altig et al. (2011) and own calculations.

* By an aggregate technology shock in Altig et al. (2011) we mean their neutral technology shock.
Figure 2: Impulse responses of consumption in the baseline economy

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