Fiscal Policy Under Rational Inattention

Piotr Dworczak*

August, 2011

Working paper

Abstract

We study fiscal policy in a DSGE model with rational inattention, thereby filling an important gap in the existing literature. To this end, we extend the fiscal sector in the model of Mackowiak and Wiederholt (2010) and propose a new approach to solving the attention problem of agents, a concept that we refer to as dynamic allocation of attention. This otherwise frictionless model can account for the puzzling finding that empirical estimates based on war episodes place the value of the fiscal multiplier systematically lower than the ones based on the VAR approach. Thanks to dynamic allocation of attention, our model predicts that when fiscal shocks are small, the crowding out effect is relatively weak and the fiscal multiplier approaches one. However, when an increase in government spending is considerable, crowding out gets larger and the multiplier is significantly lower than one.

Keywords: Rational Inattention, Fiscal Policy, General Equilibrium
JEL codes: C6, D8, E6

*University of Warsaw, Warsaw School of Economics

1 Introduction

The concept of rational inattention was introduced by Christopher Sims in his two papers Sims (1998) and Sims (2003). The basic idea behind this theory is that processing information is costly. In information-rich environments, like modern economies, the agents may find it unreasonable to use all available sources of information. They would rather focus on selected sources that provide sufficiently accurate information while limiting the amount of resources (for example time and energy) consumed by the activity of analyzing data and making decisions. More formally, we can think of agents as having a limited amount of attention that can be allocated to different activities. Processing information uses up the scarce attention resource.

As an example consider the price-setting decision of a decision maker in a firm. First of all, she must decide how much attention to allocate to that particular type of decision (other activities may include deciding about the technology of production, choosing a marketing strategy or dealing with financial issues). Secondly, she must determine how much attention to allocate to different sources of information about the optimal price. For example she may decide to learn more about the prices set by the competitors and ignore information concerning monetary policy and inflation projections. The decision maker faces a trade-off: paying more attention to a given source of information improves the quality of the decision but requires information flow that uses up valuable stock of attention. It is intuitively clear that the allocation of attention will be determined by the cost of information flow on the one hand and the scale of possible losses caused by suboptimal decisions on the other. In short, rational inattention implies that the agents choose optimally the structure of uncertainty about the economy.

We exploit this idea and develop a dynamic stochastic general equilibrium model where agents (firms and households) are rationally inattentive. Using the notion of entropy and mutual information from information theory we introduce an explicit constraint on information flow to the optimization problems of agents. We show that when processing information is costly, the resulting equilibrium differs significantly from the equilibrium that would occur under perfect information usually assumed in DSGE models. In particular, an otherwise frictionless model with rational inattention can account for many observed economic phenomena, like price and wage stickiness, humped-shaped response of some variables to shocks or different speed at which firms adjust their prices depending on the type of shock that triggers the price change.

The literature on rational inattention is still modest but expanding rapidly. The concept proved very fruitful in areas like the theory of consumer choice (Luo, 2008), monetary economics (Sims, 2005), financial economics (Batchuluun *et al.*, 2007, Luo, 2006, Mondria, 2010, Peng and Xiong, 2005), models of price setting (Matejka, 2010, Mackowiak and Wiederholt, 2007) and general equilibrium models (Luo and Young, 2009, Mackowiak and Wiederholt, 2010, Martins and Sinigaglia, 2009, Paciello, 2010). This paper belongs to that last strand of research.

There are two approaches to modelling rational inattention in DSGE models. The first (tapped by Luo and Young, 2009 and Martins and Sinigaglia, 2009) uses the social planner's solution. While this approach provides a comparatively easy method of solving the model, it fails to find justification in the form of the second welfare theorem. It also omits prices which seem to be vital in models where both agent-specific and aggregate information plays a crucial role. The second approach (Mackowiak and Wiederholt, 2010, Paciello, 2010 and also this paper) attempts to find the decentralized equilibrium and introduces information-processing constraints on the level of individual agents. Unfortunately, at the current state of our knowledge, we are doomed to computationally complex numerical methods to search for the equilibrium fixed point when we use this approach. The proof of convergence of the method or proposing a more efficient algorithm is a big challenge for future research.

Our main contribution to the rational inattention literature is the analysis of the effects of fiscal policy. Whereas monetary policy was meticulously examined by Mackowiak and Wiederholt (2010) and Paciello (2010), fiscal policy has been omitted. We introduce distortionary taxation and a government spending shock to the model that can be seen as a direct extension of Mackowiak and Wiederholt (2010) and examine the way rational inattention influences the behavior of main economic variables following a temporary increase in government spending. We show that the model can account for some of the puzzling observations about fiscal policy that other models could have a hard time to explain. For example, Bilbiie *et al.* (2006) report that US fiscal policy seems to have had stronger effects before 1980 and weaker afterwards. If we agree that the costs of processing information has fallen significantly as a result of technological progress (widespread usage of computers and the Internet) then our model gives an elegant explanation of this observation.

We find that rational inattention generally decreases crowding-out and thus makes the contemporaneous multiplier bigger. The extent to which crowding-out is reduced in comparison to the perfect information case depends on the equilibrium allocation of attention. When agents allocate much attention to observing the fiscal shock, its effectiveness is reduced. On the other hand, when the attention level is low the effects of increased government spending are stronger.

The empirical literature on the effects of fiscal expansions is vast and diversified. As far as fiscal multipliers are concerned even when ignoring extreme views about fiscal policy, we can find a whole span of values in the literature. The contemporaneous multiplier varies from 0.4-0.5 (Barro and Redlick, 2009) through 0.47-0.55 (Hall, 2009), 0.6-1.1 (Ramey, 2011) and 0.96 (Blanchard and Perotti, 2002) to values slightly above 1 (Fatas and Mihov, 2001, Rotemberg and Woodford, 1992). Most empirical studies fall into one of the two categories. Some authors (Ramey and Shapiro, 1998, Edelberg *et al.*, 1999, Burnside *et al.*, 2000, Barro and Redlick, 2009, Hall, 2009) concentrate on military spending and war episodes to quantify the effects of increased government spending. The second approach uses the Vector Autoregression approach (VAR) (see Mountford and Uhlig, 2008, Blanchard and Perotti, 2002, Fatas and Mihov, 2001). One pattern to be observed is that the estimated multipliers are generally smaller (usually below one) in the first strand of literature and bigger in the second (near or slightly above one).

We propose an extension of the model that would explain why such a pattern might take place. Concentrating on war episodes and defense spending means the fiscal shocks being studied are big, whereas the VAR approach takes into account whole time series, which mostly cover periods in which nothing particularly dramatic happens with government spending. In the model, the allocation of attention depends directly on the perceived stability of fiscal policy (the variance of the fiscal shock). When a big increase in government spending takes place agents update their estimate of the variance of the fiscal shock and increase attention capacity which leads to higher crowding-out. That may be why empirical studies based on fiscal expansions associated with wars come up with lower estimates of the multiplier. The prediction of the model can be summarized by a statement that the size of the shock matters. This is not the case for a majority of DSGE models considered in the literature.

The extension we propose could be succinctly described as dynamic allocation of attention. We argue that the assumption made in the existing rational inattention literature (for example Mackowiak and Wiederholt, 2007, Mackowiak and Wiederholt, 2010 and Paciello, 2010) that the allocation of attention is constant over time is unsatisfactory. For example it is implausible that the agents do not change their allocation of attention in the face of a huge economic turbulence, e.g. the financial crisis that happened in 2007. To capture this idea we solve the attention problem with agents allowed to choose their level of attention in every period adjusting information flows to current economic conditions.

Note that the conclusions outlined above have serious implications for fiscal policy makers. If this is really the case that big fiscal expansions are associated with lower fiscal multipliers, then this should be taken into account when designing stimulus packages like the ones tapped during the financial crisis. The mechanism of dynamic allocation of attention can be seen as yet another argument against fiscal intervention as a remedy to economic crisises.

We also study the interactions between fiscal and monetary policy under rational inattention. Using a standard monetary policy rule we show that when the nominal interest rate reacts weakly to inflation and output gap, the effects of rational inattention on the side of firms and households reinforce each other in limiting the crowding-out of consumption. However, when monetary policy becomes more aggressive, rational inattention on the side of firms works in the other way - it reduces the effectiveness of fiscal policy. The ultimate effect depends on the attention allocation of firms and households.

The remainder of this paper is organized as follows. Section 2 presents the model with perfect information (which will serve as a benchmark) and the model with rational inattention. Section 3 presents the main results. In section 4 we introduce the concept of dynamic allocation of attention. Section 5 concludes.

2 The model

2.1 Perfect information

In this chapter we present the assumptions of the model when all agents have perfect information. The economy consists of J households (indexed by $j = 1, 2, ..., J$) I firms (indexed by $(i = 1, 2, ..., I)$, a fiscal authority and a monetary authority.

We use the same notation as Mackowiak and Wiederholt (2010) to simplify comparing the models. In particular, we use lower case to indicate a variable in log-deviation from its steady state and upper case without time subscript to indicate the steady state value of a variable.

2.1.1 Households

Every household maximizes expected discounted utility from consumption and leisure subject to a budget constraint and taking as given the demand for labor.

$$\max_{\{C_{1jt},\dots,C_{Ijt},W_{jt},B_{jt}\}_{t=0}^{\infty}} E_{j,-1}\left\{\sum_{t=0}^{\infty}\beta^t\left(\frac{C_{jt}^{1-\gamma}-1}{1-\gamma}-\varphi\frac{L_{jt}^{1+\psi}}{1+\psi}\right)\right\} \tag{2.1}$$

subject to

$$\sum_{i=1}^{I} P_{it}C_{ijt} + B_{jt} = R_{t-1}B_{jt-1} + (1-\tau_t^w)W_{jt}L_{jt} + \frac{D_t}{J} - \frac{T_t}{J} \tag{2.2}$$

$$L_{jt} = d_w(W_{jt}, W_t, L_t) \tag{2.3}$$

C_{jt} is a Dixit-Stiglitz consumption aggregator[1] defined as

$$C_{jt} = \left(\sum_{i=1}^{I} C_{ijt}^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}} \tag{2.4}$$

The remaining variables are defined as follows: L_{jt} denotes labor supplied by household j, P_{it} - price of the good produced by firm i, C_{ijt} - consumption of good i by household j, B_{jt} - nominal bond holdings of household j, R_t - nominal interest rate, W_{jt} - wage rate set by household j, W_t - aggregate wage rate, τ_t^w- tax rate levied on labor income, D_t - the sum of profits of all firms, T_t - total lump-sum taxes imposed by the fiscal authority, L_t - aggregate labor supply. The parameter $0 < \beta < 1$ is the discount rate, $\theta > 1$ - the elasticity of substitution between different consumption goods, $\gamma > 0$ - the inverse of the intertemporal elasticity of substitution of consumption and $\psi > 0$ and $\varphi > 0$ affect disutility from labor.

We adopt a time convention that the time$-t$ bonds are purchased at time t and yield interest according to interest rate at time t. We assume that households can be indebted up to a certain limit, $\forall t\ B_{jt} > \bar{B}$ for some fixed $\bar{B} \in \mathbb{R}$ to exclude Ponzi schemes and that all households have the same initial bond holdings, $B_{j,-1} = 0$.

The profits of firms and lump-sum taxes are assumed to be divided equally among households. We use the function $d_w(W_{jt}, W_t, L_t)$ to denote the demand function for labor which will be derived endogenously from the optimizing behavior of firms and that is taken as given by wage setters in households. Finally, $E_{j,-1}$ is the expectation operator for household j conditioned on information available in period -1. We assume that households have rational expectations.

In every period the household sets the wage rate W_{jt}, at which it commits itself to supply the demanded amount of labor, and chooses the optimal consumption mix $(C_{1jt}, C_{2jt}, ..., C_{Ijt})$.

[1] Later we refer to C_{jt} as composite consumption.

2.1.2 Firms

Firms maximize expected discounted profits

$$\max_{\{L_{i1t},...,L_{iJt},P_{it}\}_{t=0}^{\infty}} E_{i,-1}\left\{\sum_{t=0}^{\infty} Q_t\left(P_{it}Y_{it}-\sum_{j=1}^{J} W_{jt}L_{ijt}\right)\right\} \tag{2.5}$$

Y_{it} denotes the production of firm i, L_{ijt} - labor supplied by household j to firm i, Q_t is a stochastic discount factor and $E_{i,-1}$ is the expectation operator of firm i conditioned on information available at time -1. The production function takes the form

$$Y_{it} = e^{a_t}e^{a_{it}}L_{it}^{\alpha} \tag{2.6}$$

where

$$L_{it} = \sum_{j=1}^{J}\left(L_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \tag{2.7}$$

is the Dixit-Stiglitz aggregator of labor inputs, a_t is the stochastic aggregate technology index and a_{it} is the stochastic idiosyncratic technology index, both expressed in log-deviations from the steady state in which they equal 1. α is the output elasticity of composite labor and $\eta > 1$ is the elasticity of substitution between different types of labor.

The demand for the firms' production comes from households and the government

$$Y_{it} = G_{it} + C_{it} \tag{2.8}$$

$$C_{it} = d_c(P_{it}, P_t, C_t) \tag{2.9}$$

$$G_{it} = d_g(P_{it}, P_t, G_t) \tag{2.10}$$

The function $d_c(P_{it}, P_t, C_t)$ is determined endogenously from households' optimization and the function $d_g(P_{it}, P_t, G_t)$ will be described in the next paragraph, P_t denotes the aggregate price level.

Following Mackowiak and Wiederholt (2010) we assume that the stochastic discount factor takes the form

$$Q_t = \beta^t\Lambda(C_{1t}, ..., C_{Jt})\frac{1}{P_t} \tag{2.11}$$

for some function Λ that is symmetric in its arguments and twice continuously differentiable with the property that in the steady state

$$\Lambda(C_{1t}, ..., C_{Jt}) = C_j^{-\gamma} \tag{2.12}$$

In every period the firm chooses the optimal factor mix (combination of labor types) $(L_{i1t}, L_{i2t}, ..., L_{iJt})$ and sets the price P_{it} of its good i at which it supplies the demanded quantity.

2.1.3 Fiscal and monetary authority

The budget constraint of fiscal authority reads

$$R_{t-1}B_{t-1} + \sum_{i=1}^{I} P_{it}G_{it} = B_t + \tau_t^w\left(\sum_{j=1}^{J} W_{jt}L_{jt}\right) + T_t \tag{2.13}$$

That is, in each period government spending (consisting of purchases of consumption goods from firms and the payment made to time-$(t-1)$ bond holders) must equal the revenues from issuing new bonds and collecting labor and lump-sum taxes.

We define budget surplus as the relation of tax revenues to consumption expenditures

$$S_t = \frac{\tau_t^w(\sum_{j=1}^J W_{jt}L_{jt}) + T_t}{\sum_{i=1}^I P_{it}G_{it}} \tag{2.14}$$

We assume that fiscal policy follows the following policy rule[2]

$$S_t = S^\star + \omega\left(\frac{Y_t - Y}{Y}\right) \tag{2.15}$$

The rule states that the actual surplus is the steady state surplus $S^\star$ plus the cyclical component that depends on the output gap (Y denotes the value of aggregate production in the steady state). The parameter ω governs the responsiveness of fiscal policy and the approach of fiscal authorities to output fluctuation ($\omega > 0$ implies antycyclical policy).

We assume that there is a constant share of labor tax in total tax revenues, specifically in each period we have

$$\frac{1-\theta_\tau}{\theta_\tau} = \frac{T_t}{\tau_t^w(\sum_{j=1}^J W_{jt}L_{jt})} \tag{2.16}$$

for some fixed parameter $\theta_\tau \in [0,1]$. We define

$$B_t = \sum_{j=1}^J B_{jt} \tag{2.17}$$

and

$$G_t = \left(\sum_{i=1}^I G_{it}^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}} \tag{2.18}$$

We assume that the demand for the firm i's good is given by

$$G_{it} = d_g(P_{it}, P_t, G_t) = \left(\frac{P_{it}}{P_t}\right)^{-\theta} G_t \tag{2.19}$$

We make this assumption because it implies that government demand has the same form as the demand of households - this makes aggregation easier.

The total government consumption spending is determined by the following stochastic process

$$\frac{G_t}{G} = \left(\frac{G_{t-1}}{G}\right)^{\rho^G} e^{\varepsilon_t^G} \tag{2.20}$$

where G is the steady state government consumption spending assumed a fixed fraction of aggregate output, $G = \theta_g Y$.

Monetary authority sets the nominal interest rate according to a standard monetary policy rule

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\rho_R}\left[\left(\frac{\Pi_t}{\Pi}\right)^{\phi_\pi}\left(\frac{Y_t}{Y}\right)^{\phi_y}\right]^{1-\rho_R} \tag{2.21}$$

$\Pi_t = \frac{P_t}{P_{t-1}}$ denotes gross inflation rate and the parameters ϕ_π, ϕ_y and ρ_R govern the persistence and responsiveness of the interest rate to aggregate conditions. We assume that monetary policy is deterministic in the sense that there is no stochastic exogenous shock to the interest rate set by the monetary authority, as the focus of the analysis will be on fiscal policy[3].

[2]See Fueki *et al.* (2011) for comparison

[3]Monetary policy shocks have been studied in Mackowiak and Wiederholt (2010) and Paciello (2010)

2.1.4 Aggregation

First, we introduce some new notation

$$\hat{P}_{it} = \frac{P_{it}}{P_t} \tag{2.22}$$

$$\hat{W}_{jt} = \frac{W_{jt}}{W_t} \tag{2.23}$$

We also define aggregate price level and aggregate wage rate as

$$P_t = \left(\sum_{i=1}^{I} P_{it}^{1-\theta} \right)^{\frac{1}{1-\theta}} \tag{2.24}$$

$$W_t = \left(\sum_{j=1}^{J} W_{jt}^{1-\eta} \right)^{\frac{1}{1-\eta}} \tag{2.25}$$

Then we have

$$L_t = \sum_{i=1}^{I} L_{it} = \sum_{j=1}^{J} \hat{W}_{jt} L_{jt} \tag{2.26}$$

$$C_t = \sum_{i=1}^{I} \hat{P}_{it} C_{it} = \sum_{j=1}^{J} C_{jt} \tag{2.27}$$

$$G_t = \sum_{i=1}^{I} \hat{P}_{it} G_{it} \tag{2.28}$$

It is easy to verify that equation 2.28 is compatible with the assumption 2.19.

Finally we can define aggregate output as

$$Y_t = \sum_{i=1}^{I} \hat{P}_{it} Y_{it} = C_t + G_t \tag{2.29}$$

2.1.5 Shocks

There are three types of shocks in the economy: an aggregate technological shock ε_t^A, idiosyncratic technological shocks ε_{it}^I, for $i = 1, 2, ..., I$ and a fiscal shock ε_t^G. We assume that the aggregate technological index follows an AR(1) process

$$a_t = \rho_A a_{t-1} + \varepsilon_t^A \tag{2.30}$$

Similarly, for each $i = 1, 2, ..., I$ we have

$$a_{it} = \rho_I a_{it-1} + \varepsilon_{it}^I \tag{2.31}$$

The processes $\{\varepsilon_t^A\}$, $\{\varepsilon_{it}^I\}_{i=1}^I$, $\{\varepsilon_t^G\}$ are assumed to follow Gaussian white noise processes with variances σ_A^2, σ_I^2 and σ_G^2, respectively, and are independent of each other. Since we only have finitely many firms and households idiosyncratic shocks can, despite independence, affect aggregate conditions. However, we ignore this effect since when I and J are very large its impact is negligible. We simply assume that if a set of variables $\{x_{it}\}_{i=1}^I$ (or $\{x_{jt}\}_{j=1}^J$) is independent and every variable has zero mean then for every $t \geq 0$

$$\sum_{i=1}^{I} x_{it} = 0 \quad \left(\sum_{j=1}^{J} x_{jt} = 0 \right) \tag{2.32}$$

2.1.6 Equilibrium

We define the equilibrium of the model under perfect information.

Definition. The equilibrium of the model with perfect information is the allocation $\{\{C^\star_{ijt}\}^{I,J}_{i,j=1}, \{L^\star_{ijt}\}^{I,J}_{i,j=1}, \{B^\star_{jt}\}^{J}_{j=1}, \{G^\star_{it}\}^{I}_{i=1} D^\star_t, T^\star_t, \tau^{w\star}_t\}^\infty_{t=0}$, prices $\{\{W^\star_{jt}\}^J_{j=1}, \{P^\star_{it}\}^I_{i=1}, R^\star_t\}^\infty_{t=0}$ and functions $d^\star_c(P_{it}, P_t, C_t)$, $d^\star_w(W_{jt}, W_t, L_t)$, $d^\star_g(P_{it}, P_t, G_t)$ such that

1. For each j, given prices $\{P^\star_{it}\}^I_{i=1}$, taxes $T^\star_t$ and $\tau^{w\star}_t$, firms' profits $D^\star_t$, interest rate $R^\star_t$, aggregate wage rate $W^\star_t$, aggregate labor supply $L^\star_t$, initial bond holdings $B^\star_{j,-1}$ and labor demand function $d^\star_w(W_{jt}, W_t, L_t)$, $\{C_{1jt}, ..., C_{Ijt}, W_{jt}, B_{jt}\}^\infty_{t=0}$ solves the problem of the household: 2.1 subject to 2.2, 2.3 and 2.4,

2. For each i, given wage rates $\{W^\star_{jt}\}^J_{j=1}$, aggregate consumption $C^\star_t$, aggregate price level $P^\star_t$ and demand functions $d^\star_c(P_{it}, P_t, C_t)$ and $d^\star_g(P_{it}, P_t, G_t)$ $\{L_{i1t}, ..., L_{iJt}, P_{it}\}^\infty_{t=0}$ solves the problem of the firm: 2.5 subject to 2.6 - 2.11,

3. $\{\{B^\star_{jt}\}^J_{j=1}, \{G^\star_{it}\}^I_{i=1}, T^\star_t, \tau^{w\star}_t, R^\star_t\}^\infty_{t=0}$ and $d^\star_g(P_{it}, P_t, G_t)$ satisfy equations 2.13 - 2.21,

4. Function $d^\star_c(P_{it}, P_t, C_t)$ can be obtained from the optimal choices of households in point 1 and function $d^\star_w(W_{jt}, W_t, L_t)$ can be obtained from the optimal choices of firms in point 2 after using definitions 2.22 - 2.28,

5. Markets clear: $\forall j \, L^\star_{jt} = d^\star_w(W^\star_{jt}, W^\star_t, L^\star_t)$, $\forall i \, C^\star_{it} = d^\star_c(P^\star_{it}, P^\star_t, C^\star_t)$

The existence and uniqueness of the equilibrium defined above is guaranteed by standard Blanchard-Kahn[4] conditions. We always check whether these conditions are satisfied for given parameter values.

2.2 Imperfect information

We now introduce rational inattention to the model. Following the literature on rational inattention we use the notion of entropy and mutual information to measure information flow[5]. We will use the fact that for processes $\{X_t\}^\infty_{t=0}$, $\{Z_t\}^\infty_{t=0}$ that are stationary and Gaussian the following formula describes how much information about the process $\{X_t\}^\infty_{t=0}$ is contained in the process $\{Z_t\}^\infty_{t=0}$

$$\mathcal{I}(\{X_t\}, \{Z_t\}) = \lim_{T\to\infty} \frac{1}{T} I(X_0, X_1, ..., X_{T-1}; Z_0, Z_1, ..., Z_{T-1}) \tag{2.33}$$

where

$$I(X_0, X_1, ..., X_{T-1}; Z_0, Z_1, ..., Z_{T-1}) = \frac{1}{2}\log_2\left(\frac{\det \Omega_X}{\det \Omega_{X|Z}}\right) \tag{2.34}$$

where Ω_X is the covariance matrix of the vector $(X_0, X_1, ..., X_{T-1})$ and $\Omega_{X|Z}$ is the conditional covariance matrix of the vector $(X_0, X_1, ..., X_{T-1})$ given vector $(Z_0, Z_1, ..., Z_{T-1})$.

Following Mackowiak, Wiederholt Mackowiak and Wiederholt (2010) we model rational inattentiveness of agents as the existence of an additional constraint of the form $\mathcal{I}(\{X_t\}, \{Z_t\}) \le \kappa$, where κ is some positive number. As the processes $\{X_t\}^\infty_{t=0}, \{Z_t\}^\infty_{t=0}$ we take the optimal and actual decision made by an agent which means that households and firms have to take into account the informational content of the decisions they make.

In the following sections we describe the modifications to the optimization problems of the household and the firm introduced in section 2.1. We present the problems in a log-quadratic setting that is derived from the original problem by log-deviating all variables and taking the second-order approximation of the objective function. The derivation can be found in the Technical Appendix. We remind that variables in lower case denote log deviation from the steady state.

[4]see Blanchard and Kahn (1980)

[5]See Sims (2003) or Sims (2010) for details. For a more systematic introduction to information theory see Cover and Thomas (2006)

2.2.1 The problem of the firm

In the model with rational inattention the firm solves the following optimization problem

$$\max_{\kappa,\{B_i(L),C_i(L)\}_{i=1}^{3},\tilde{\eta},\chi}\left\{\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}(x_t-x_t^\star)'H(x_t-x_t^\star)\right]-\frac{\mu}{1-\beta}\kappa\right\} \tag{2.35}$$

where

$$x_t-x_t^\star=\begin{pmatrix} p_{it}\\ \hat{l}_{i1t}\\ \vdots\\ \hat{l}_{i(J-1)t}\end{pmatrix}-\begin{pmatrix} p_{it}^\star\\ \hat{l}_{i1t}^\star\\ \vdots\\ \hat{l}_{i(J-1)t}^\star\end{pmatrix} \tag{2.36}$$

$$H\in\mathbb{R}^{(J+1)\times(J+1)}\ \text{(negative-definite)}$$

$$p_{it}^\star=\underbrace{A_1(L)\varepsilon_t^A}_{p_{it}^{A\star}}+\underbrace{A_2(L)\varepsilon_{it}^I}_{p_{it}^{I\star}}+\underbrace{A_3(L)\varepsilon_t^G}_{p_{it}^{G\star}} \tag{2.37}$$

$$\hat{l}_{ijt}^\star=-\eta\hat{w}_{jt}\quad j=1,2,...,J-1 \tag{2.38}$$

$$p_{it}=\underbrace{B_1(L)\varepsilon_t^A+C_1(L)\nu_{it}^A}_{p_{it}^{A}}+\underbrace{B_2(L)\varepsilon_{it}^I+C_2(L)\nu_{it}^I}_{p_{it}^{I}}+\underbrace{B_3(L)\varepsilon_t^G+C_3(L)\nu_{it}^G}_{p_{it}^{G}} \tag{2.39}$$

$$\hat{l}_{ijt}=-\tilde{\eta}\left(\hat{w}_{jt}+\chi\nu_{ijt}^L\right)\quad j=1,2,...,J-1 \tag{2.40}$$

subject to the information processing constraint

$$\mathcal{I}\left(\left\{p_{it}^\star,\hat{l}_{i1t}^\star,...,\hat{l}_{i(J-1)t}^\star\right\};\left\{p_{it},\hat{l}_{i1t},...,\hat{l}_{i(J-1)t}\right\}\right)\le\kappa \tag{2.41}$$

At first the form of the above problem may be a little puzzling, so the remainder of this section is devoted to its discussion. The reader is referred to the Technical Appendix for further details.

First, the first part of the expression 2.35 is the expected discounted difference between the value of the objective function when the choice variables are given by x_t (vector x_t gathers all choice variables expressed in log-deviations from the steady state) and the value of the objective function when x_t is chosen optimally, that is, under perfect information (the optimal choice is indicated by a star, e.g. $x_t^\star$). Thus, we can view the problem as minimizing the expected loss arising from suboptimal choice of x_t which is equivalent to maximizing the original objective function. Second, it can be showed that the original constrained problem can be transformed into an unconstrained problem (apart from the information processing constraint) where the choice variables are $\hat{p}_{it}$, the relative price defined as $\hat{p}_{it}=p_{it}-p_t$ (or $\hat{P}_{it}=\frac{P_{it}}{P_t}$) and shares of different types of labor in composite labor, $\hat{l}_{ijt}=l_{ijt}-l_{it}$ (or $\hat{L}_{ijt}=\frac{L_{ijt}}{L_{it}}$). Third, equations 2.37 and 2.38 characterize the optimal decision, or more precisely, the solution of the firms's optimization problem that would arise if the firm had perfect information. $A_1(L)$, $A_2(L)$, $A_3(L)$ denote infinite-order lag polynomials with absolutely summable coefficients. The optimal price at time t can be represented as a linear function (or moving average process) of all the shocks up to and including time t[6]. The response of the price to shocks is divided into three separate components. $p_{it}^{A\star}$ is the response to aggregate technological shocks, $p_{it}^{I\star}$ - the response to idiosyncratic technological shocks and $p_{it}^{G\star}$- to fiscal shocks. The optimal share of the $j-th$ type of labor in composite labor is optimally a linear function of the relative wage set by household j, that is, $\hat{w}_{jt}=w_{jt}-w_t$. In particular it does not depend on aggregate conditions. Fourth, the actual decision is given by equations 2.39 and 2.40. The stochastic variables ν_{it}^A, ν_{it}^I, ν_{it}^G, $\left\{\nu_{ijt}^L\right\}_{j=1}^{J-1}$ are assumed to follow white noise Gaussian processes with unit variance and are independent of each other (also independent across firms). These variables represent noise in the actual decisions. The actual decision can thus

[6]For example if $A=(a_0,a_1,a_2,...)$ then $A(L)\varepsilon_t=\sum_{i=0}^{\infty}a_i\varepsilon_{t-i}$

be seen as a combination of a pure signal which gives information about the optimal decision and noise which contaminates the decision and decreases its informational content. The firm determines the value of the choice variables by choosing the coefficients of infinite-order lag polynomials $B_1(L)$, $B_2(L)$, $B_3(L)$ and $C_1(L)$, $C_2(L)$, $C_3(L)$ and choosing the parameters $\tilde{\eta}$ and χ. For example, if there was no information processing constraint, the firm would choose $B_1(L) = A_1(L)$, $B_2(L) = A_2(L)$, $B_3(L) = A_3(L)$, $C_1(L) = C_2(L) = C_3(L) = 0$, $\tilde{\eta} = \eta$, $\chi = 0$ implying that the actual decision would equal the optimal decision in every period. The reason we do not let the firm choose arbitrary lag polynomials in the case of labor mix decision is simple: by assuming the structural form 2.40 we are able to obtain an analytically tractable expression for the demand for labor on and off the equilibrium path which enables us to find the general equilibrium. Still, by choosing $\tilde{\eta}$ and χ the firm can regulate the precision and information content of the actual decision concerning factor mix. Fifth, notice that given the form of the actual and optimal decision we can compute the value of the expected discounted loss in profits in the objective 2.35 as it depends only on the coefficients of the lag polynomials (or $\tilde{\eta}$ and χ) and the variance of the shocks, σ_A^2, σ_G^2 or σ_I^2. However, we have to make an assumption concerning the realizations of shocks before period 0. It seems natural to assume the the shocks were zero and that the agents are aware of this fact, so that past realizations of shocks do not influence the value of the objective 2.35. Finally, the inequality 2.41 constrains the informational content of the actual decision. Roughly speaking this condition says that the actual decision cannot provide too much information about the optimal decision or, put differently, the closer the agent wants to get to the optimal decision, the more attention she needs to allocate. The level of attention is measured by κ. Instead of treating κ as a given parameter we let the firm choose its total attention capacity. Allocating more attention is costly: it decreases expected profits proportionally to a parameter μ which denotes the marginal cost of processing information. Thus the firm faces the following trade-off: allocating more attention decreases profits by increasing the costly information flow, but decreases the loss in profits as the actual decision gets closer to the optimal decision. One important observation needs to be made here. By assuming independence between noise components ν_{it} and thanks to the additive (across shocks) representation 2.37 and 2.39, we can divide information flow in the constraint 2.41 into three independent parts, $\kappa = \kappa_A + \kappa_I + \kappa_G$, which correspond to information flows connected with aggregate technological shocks, idiosyncratic technological shock and fiscal shock, respectively[7]. It means that we have implicitly assumed that attending aggregate technological conditions, idiosyncratic technological conditions and fiscal conditions are separate activities, each requiring a separate information flow. This simplifying assumption can be justified by the interpretation that rational inattention captures information processing constraints rather then information scarcity[8].

2.2.2 The problem of the household

We make a simplifying assumption regarding the decision-making process in households. We assume that every household consists of a wage-setter and a consumer. The wage-setter sets the wage rate W_{jt}, whereas the consumer takes the income of the household as given and decides how to spend it on bond purchases B_{jt} and consumption goods $(C_{1jt}, ..., C_{Ijt})$. In particular, the two decisionmakers have separate information constraints and information flows devoted to making one decision do not influence the other. Once again this assumption can be justified by the interpretation that rationally inattentive agents face information processing constraints rather then lack of information. Therefore we can treat those two decisions as separate although they require information about the same aggregate variables[9].

After taking the log-quadratic approximation of the objective function the problem of the wage-setter can be expressed as

[7] We use the additivity property of mutual information, see Cover and Thomas (2006) for details and proof

[8] See Sims (2010) for a further discussion of this problem in the rational inattention literature

[9] A similar idea is exploited by Paciello (2010)

$$\max_{\kappa,\{B_i(L),C_i(L)\}_{i=1}^2} \left\{ \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[\frac{1}{2}(x_t^w - x_t^\star)' H_0 (x_t^w - x_t^\star) + (x_t^w - x_t^\star)' H_1 (x_{t+1}^w - x_{t+1}^\star) \right] - \frac{\lambda}{1-\beta}\kappa \right\} \tag{2.42}$$

where

$$x_t^w - x_t^\star = \begin{pmatrix} \tilde{b}_{jt}^\star \\ \tilde{w}_{jt} \\ \hat{c}_{1jt}^\star \\ \vdots \\ \hat{c}_{(I-1)jt}^\star \end{pmatrix} - \begin{pmatrix} \tilde{b}_{jt}^\star \\ \tilde{w}_{jt}^\star \\ \hat{c}_{1jt}^\star \\ \vdots \\ \hat{c}_{(I-1)jt}^\star \end{pmatrix} \tag{2.43}$$

$$H_0,\ H_1 \in \mathbb{R}^{(I+2)\times(I+2)}$$

$$\tilde{w}_{jt}^\star = \underbrace{A_1(L)\varepsilon_t^A}_{\tilde{w}_{jt}^{A\star}} + \underbrace{A_2(L)\varepsilon_t^G}_{\tilde{w}_{jt}^{G\star}} \tag{2.44}$$

$$\tilde{w}_{jt} = \underbrace{B_1(L)\varepsilon_t^A + C_1(L)v_{jt}^A}_{\tilde{w}_{jt}^A} + \underbrace{B_2(L)\varepsilon_t^G + C_2(L)v_{jt}^G}_{\tilde{w}_{jt}^G} \tag{2.45}$$

subject to the information processing constraint

$$\mathcal{I}\left(\left\{\tilde{w}_{jt}^\star\right\};\left\{\tilde{w}_{jt}\right\}\right) \leq \kappa \tag{2.46}$$

The form of this problem is analogous to the form of the firm's problem. The objective function 2.42 is the expected discounted loss in welfare due to suboptimal choice of wage rate net of information processisng costs. The parameter λ is the marginal welfare cost of information. $x_t^\star$ denotes the solution of the problem under perfect information (more precisely: the solution of the wage setter's problem that would arise if the wage setter had perect information). We introduced the additional notation (tilde) to denote real variables when the original variable was expressed in nominal terms. Specifically, $\tilde{b}_{jt} = b_{jt} - p_t$ and $\tilde{w}_{jt} = w_{jt} - p_t$. The optimal real wage at time t is a linear function of the realization of shocks up to and including time t according to 2.44. The actual wage rate is given by 2.45 and is determined by the coefficients of the infinite-order lag polynomial chosen by the wage-setter. The variables v_{jt}^A, v_{jt}^G are assumed to follow a Gaussian white noise process with unit variance and are independent of each other and all other stochastic processes in the economy. The information flow in constraint 2.46 depends only on the actual and optimal wage decision. One thing needs to be underscored here. When we state the problem of the wage-setter we assume that the decisions made by the consumer are optimal. Hence all but one elements of the vector $x_t^w - x_t^\star$ are identically zero. In practice this means that we ignore possible interdependency between the wage-setting and spending decision. Later we check whether the results that we obtain are robust to this assumption. It turns out that when we replace the optimal decision of the consumer in the wage-setting problem by the actual decision of the consumer, the results remain unchanged[10].

The problem of the consumer reads

$$\max_{\kappa,\{B_i(L),C_i(L)\}_{i=1}^2,\tilde{\theta},\vartheta} \left\{ \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[\frac{1}{2}(x_t - x_t^\star)' H_0 (x_t - x_t^\star) + (x_t - x_t^\star)' H_1 (x_{t+1} - x_{t+1}^\star) \right] - \frac{\lambda}{1-\beta}\kappa \right\} \tag{2.47}$$

[10]We do this by plugging a guess of the actual spending decision made by the consumer into the wage-setting problem in every iteration of the solution procedure. At the equilibrium fixed point the guess agrees with the actual decision of the consumer.

where

$$x_t - x_t^\star = \begin{pmatrix} \tilde{b}_{jt} \\ \tilde{w}_{jt}^\dagger \\ \hat{c}_{1jt} \\ \vdots \\ \hat{c}_{(I-1)jt} \end{pmatrix} - \begin{pmatrix} \tilde{b}_{jt}^\star \\ \tilde{w}_{jt}^\star \\ \hat{c}_{1jt}^\star \\ \vdots \\ \hat{c}_{(I-1)jt}^\star \end{pmatrix} \tag{2.48}$$

$$H_0,\ H_1 \in \mathbb{R}^{(I+2)\times(I+2)}$$

$$b_{jt} - b_{jt}^\star = -\sum_{l=0}^{t} \left(\frac{1}{\beta}\right)^l \frac{1}{\omega_B} \left[\left(c_{j,t-l} - c_{j,t-l}^\star\right) + \tilde{\eta}\omega_W \left(w_{j,t-l}^\dagger - w_{j,t-l}^\star\right)\right] \tag{2.49}$$

$$c_{jt}^\star = \underbrace{A_1(L)\varepsilon_t^A}_{c_{jt}^{A\star}} + \underbrace{A_2(L)\varepsilon_t^G}_{c_{jt}^{G\star}} \tag{2.50}$$

$$c_{jt} = \underbrace{B_1(L)\varepsilon_t^A + C_1(L)\xi_{jt}^A}_{c_{jt}^A} + \underbrace{B_2(L)\varepsilon_t^G + C_2(L)\xi_{jt}^G}_{c_{jt}^G} \tag{2.51}$$

$$\hat{c}_{ijt}^\star = -\theta \hat{p}_{it} \quad i = 1, 2, ..., I-1 \tag{2.52}$$

$$\hat{c}_{ijt} = -\tilde{\theta}\left(\hat{p}_{it} + \vartheta \xi_{ijt}^C\right) \quad i = 1, 2, ..., I-1 \tag{2.53}$$

subject to the information processing constraint

$$\mathcal{I}\left(\left\{c_{jt}^\star, \hat{c}_{1jt}^\star, ..., \hat{c}_{(I-1)jt}^\star\right\}; \left\{c_{jt}, \hat{c}_{1jt}, ..., \hat{c}_{(I-1)jt}\right\}\right) \leq \kappa \tag{2.54}$$

When the consumer solves her optimization problem she takes as given the wage-setting decision made by the wage-setter, dentoed by $\tilde{w}_{jt}^\dagger$. Hence $\tilde{w}_{jt}^\dagger$ is treated as an exogenous variable. We use the budget constraint to derive the equation 2.49 which (given the stochastic process for $\hat{w}_{jt}^\dagger$) establishes a one-to-one correspondence between the consumption path$\{c_{jt}\}_{t=0}^\infty$ and bond holdings path $\{b_{jt}\}_{t=0}^\infty$. In consequence we can focus on the choice of the process for composite consumption c_{jt} rather than the process for bond purchases b_{jt}. Apart from this the consumer makes decision regarding the optimal consumption mix, that is, the shares of different goods i in the composite consumption: $\hat{c}_{ijt} = c_{ijt} - c_{jt}$ for $i = 1, 2, ..., I-1$. Under perfect information the share of good i depends only on its relative price $\hat{p}_{it} = p_{it} - p_t$. For the reasons made explicit in the discussion of the firm's problem we assume that the choice of the actual share $\hat{c}_{ijt}$ is restricted to the form 2.53. As before we assume that the noise variables ξ_{jt}^A, ξ_{jt}^G, $\{\xi_{ijt}^C\}_{i=1}^{I-1}$ follow a Gaussian white noise process with unit variance and are independent of each other and all other stochastic processes in the economy.

2.2.3 Aggregation under rational inattention

A few remarks are needed to describe the necessary changes to the way we aggregate certain variables in the model with rationally inattentive agents. The need for change stems from the fact that the actual elasticities of substitution between different types of consumption goods and different types of labor are no longer given by θ and η (as in the perfect information case), but by $\tilde{\theta}$ and $\tilde{\eta}$ chosen optimally under information processing constraints by households and firms. Hence we define

$$P_t = \left(\sum_{i=1}^{I} P_{it}^{1-\tilde{\theta}}\right)^{\frac{1}{1-\tilde{\theta}}} \tag{2.55}$$

$$W_t = \left(\sum_{j=1}^{J} W_{jt}^{1-\tilde{\eta}}\right)^{\frac{1}{1-\tilde{\eta}}} \tag{2.56}$$

We also assume that $\tilde{\theta}$ replaces θ in equation 2.19. When we calibrate the model we use the data to determine the value of the actual elasticities $\tilde{\theta}$ and $\tilde{\eta}$ and then choose values for θ and η which yield $\tilde{\theta}$ and $\tilde{\eta}$ in equilibrium. This is possible since it turns out that we can solve the factor mix problem of the firm and the consumption mix problem of the household analytically independently of other components of the optimization problems.

2.2.4 Equilibrium

Definition. The equilibrium of the model with rational inattention is the allocation $\{\{c^{\dagger}_{ijt}\}^{I,J}_{i,j=1}, \{l^{\dagger}_{ijt}\}^{I,J}_{i,j=1}, \{\tilde{b}^{\dagger}_{jt}\}^{J}_{j=1}, \{g^{\dagger}_{it}\}^{I}_{i=1}\, \tilde{d}^{\dagger}_{t},\, \tilde{t}^{\dagger}_{t},\, \tau^{\dagger}_{t}\}^{\infty}_{t=0},\, \{\kappa^{\dagger}_{i}\}^{I}_{i=1},\, \{\kappa^{w\dagger}_{j},\, \kappa^{c\dagger}_{j}\}^{J}_{j=1}$,
prices $\{\{\tilde{w}^{\dagger}_{jt}\}^{J}_{j=1}, \{p^{\dagger}_{it}\}^{I}_{i=1},\, r^{\dagger}_{t}\}^{\infty}_{t=0}$
and functions $d^{\dagger}_{c}(p_{it}, p_t, c_t)$, $d^{\dagger}_{w}(\tilde{w}_{jt}, \tilde{w}_t, l_t)$, $d^{\dagger}_{g}(p_{it}, p_t, g_t)$ such that

1. For each j, given prices $\{p^{\dagger}_{it}\}^{I}_{i=1}$, taxes $\tilde{t}^{\dagger}_{t}$ and $\tau^{\dagger}_{t}$, firms' profits $\tilde{d}^{\dagger}_{t}$, interest rate $r^{\dagger}_{t}$, real aggregate wage rate $\tilde{w}^{\dagger}_{t}$, aggregate labor supply $l^{\dagger}_{t}$, initial bond holdings $\tilde{b}^{\dagger}_{j,-1}$ and labor demand function $d^{\dagger}_{w}(\tilde{w}_{jt}, \tilde{w}_t, l_t)$, there exist infinite-order lag polynomials with absolutely summable coefficients $\{B^{w}_{jk}(L), C^{w}_{jk}(L), B^{c}_{jk}(L), C^{c}_{jk}(L)\}^{2}_{k=1}$ and numbers $\tilde{\theta}_j$ and ϑ_j such that $\{\kappa^{w}_{j},\, \{B^{w}_{jk}(L), C^{w}_{jk}(L)\}^{2}_{k=1}\}$ solves the problem of the household: 2.42 - 2.46, $\{\kappa^{c}_{j},\, \{B^{c}_{jk}(L), C^{c}_{jk}(L)\}^{2}_{k=1}, \tilde{\theta}_j, \vartheta_j\}$ solves the problem 2.47 - 2.54 and
$$\tilde{w}^{\dagger}_{jt} = B^{w}_{j1}(L)\varepsilon^{A}_{t} + C^{w}_{j1}(L)v^{A}_{jt} + B^{w}_{j2}(L)\varepsilon^{G}_{t} + C^{w}_{j2}(L)v^{G}_{jt}$$
$$c^{\dagger}_{jt} = B^{c}_{j1}(L)\varepsilon^{A}_{t} + C^{c}_{j1}(L)\xi^{A}_{jt} + B^{c}_{j2}(L)\varepsilon^{G}_{t} + C^{c}_{j2}(L)\xi^{G}_{jt}$$
$$\hat{c}^{\dagger}_{ijt} = -\tilde{\theta}_j\left(\hat{p}^{\dagger}_{it} + \vartheta_j\xi^{C}_{ijt}\right),\;\; i = 1, 2, ..., I-1$$

2. For each i, given wage rates $\{\tilde{w}^{\dagger}_{jt}\}^{J}_{j=1}$, aggregate consumption $c^{\star}_{t}$, aggregate price level $p^{\star}_{t}$ and demand functions $d^{\dagger}_{c}(p_{it}, p_t, c_t)$ and $d^{\dagger}_{g}(p_{it}, p_t, g_t)$, there exist infinite-order lag polynomials with absolutely summable coefficients $\{B_{ik}(L), C_{ik}(L)\}^{3}_{k=1}$ and numbers $\tilde{\eta}_i$ and χ_i such that $\{\kappa^{\dagger}_{i}, \{B_{ik}(L), C_{ik}(L)\}^{3}_{i=1}, \tilde{\eta}_i, \chi_i\}$ solves the problem of the firm: 2.35 - 2.41 and
$$p^{\dagger}_{it} = B_{i1}(L)\varepsilon^{A}_{t} + C_{i1}(L)\nu^{A}_{it} + B_{i2}(L)\varepsilon^{I}_{it} + C_{i2}(L)\nu^{I}_{it} + B_{i3}(L)\varepsilon^{G}_{t} + C_{i3}(L)\nu^{G}_{it}$$
$$\hat{l}^{\dagger}_{ijt} = -\tilde{\eta}_i\left(\hat{w}^{\dagger}_{jt} + \chi_i\nu^{L}_{ijt}\right),\;\; j = 1, 2, ..., J-1$$

3. $\{\{\tilde{b}^{\dagger}_{jt}\}^{J}_{j=1}, \{g^{\dagger}_{it}\}^{I}_{i=1},\, \tilde{t}^{\dagger}_{t},\, \tau^{\dagger}_{t}, r^{\dagger}_{t}\}^{\infty}_{t=0}$ and $d^{\dagger}_{g}(p_{it}, p_t, g_t)$ satisfy the loglinearized equations 2.13 - 2.21,

4. Function $d^{\dagger}_{c}(p_{it}, p_t, c_t)$ can be obtained from the optimal choices of households in point 1 and function $d^{\dagger}_{w}(\tilde{w}_{jt}, \tilde{w}_t, l_t)$ can be obtained from the optimal choices of firms in point 2 after using definitions 2.22 - 2.28 adjusted by modifications described in 2.2.3,

5. Markets clear: $\forall j\; l^{\dagger}_{jt} = d^{\dagger}_{w}(\tilde{w}^{\dagger}_{jt}, \tilde{w}^{\dagger}_{t}, l^{\dagger}_{t})$, $\forall i\; c^{\dagger}_{it} = d^{\dagger}_{c}(p^{\dagger}_{it}, p^{\dagger}_{t}, c^{\dagger}_{t})$.

2.2.5 Equilibrium existence and computation

The proof of existence and uniqueness of the above defined equilibrium is beyond the scope of this work. However, we present and prove one important observation which will be used to find the equilibrium of the model numerically.

We prove in the Technical Appendix that the choice of the optimal consumption mix 2.53 and the optimal factor mix 2.40 is independent of the choice of other variables and can be derived analytically. We can thus focus on the choice of the optimal price level by the firm, optimal wage rate by the wage-setter and optimal composite consumption by the consumer and on the two aggregate shocks (we omit the idiosyncratic technological shock in further analysis). Suppose there exists a unique solution to those problems and a unique solution to the corresponding problems with perfect information. Next notice that the independence assumption (between shocks and between noise in different components of choice variables) combined with the additivity property of mutual information and linearity of the objective function in total attention capacity κ implies that we can solve the model independently for each type of shock. Consider the fiscal shock. We can define a mapping $\Phi_G : \ell^1 \times \ell^1 \times \ell^1 \to \ell^1 \times \ell^1 \times \ell^1$ which takes as an argument a triple $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$ with each element represented as an infinite absolutely summable sequence (element of the ℓ^1 space). This representation is possible since we can treat the elements of the sequence as the coefficients of the infinite-order lag polynomial and each element of $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$ as a linear function of the fiscal shocks up to and including time t as in 2.37, 2.44 and 2.50. The value of the mapping Φ_G at $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$ is determined in the following steps:

1. We treat $p_{it}^{G\star}$ as the profit-maximizing price response to fiscal shock in the firm's problem (2.37), $\tilde{w}_{jt}^{G\star}$ as the optimal wage rate response to fiscal shock in the wage-setting problem (2.44) and $c_{jt}^{G\star}$ as the optimal composite consumption response to fiscal shock in the consumer's problem (2.50). We solve those problems (under information processing constraints) and obtain a triple $(p_{it}^{G},\, \tilde{w}_{jt}^{G},\, c_{jt}^{G})$ as a solution,

2. Given the decision of firms and households $(p_{it}^{G},\, \tilde{w}_{jt}^{G},\, c_{jt}^{G})$ [11], we find the aggregate dynamics of the economy,

3. Given the aggregate dynamics we compute the profit-maximizing price for the firm $p_{it}^{G\star\star}$, the optimal wage rate for the wage-setter $\tilde{w}_{jt}^{G\star\star}$ and the optimal composite consumption for the consumer $c_{jt}^{G\star\star}$, assuming the decision makers have perfect information,

4. We assign the triple $(p_{it}^{G\star\star},\, \tilde{w}_{jt}^{G\star\star},\, c_{jt}^{G\star\star})$ represented as an element of $\ell^1 \times \ell^1 \times \ell^1$ as the value of Φ_G at $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$.

Intuitively, the mapping Φ_G transforms a guess concerning the way the agents would behave if they had perfect information into their actual optimal behavior under perfect information, but with aggreagate dynamics implied by the initial guess.

Similarly, we can define a mapping Φ_A that works analogously on a triple $(p_{it}^{A\star},\, \tilde{w}_{jt}^{A\star},\, c_{jt}^{A\star})$, that is, the component of the price, wage rate and composite consumption that responds to technological shocks. Then we define $\Phi = (\Phi_G,\, \Phi_A)$.

Having defined the mapping Φ we can state the following Proposition.

Proposition. *There exists an equilibrium of the model with rational inattention if and only if there exists a fixed point of the mapping* Φ.

Proof. " $\Longrightarrow$ " Suppose we have an equilibrium of the model with rational inattention:

$$\{\{c_{ijt}^\dagger\}_{i,j=1}^{I,J}, \{l_{ijt}^\dagger\}_{i,j=1}^{I,J}, \{\tilde{b}_{jt}^\dagger\}_{j=1}^{J}, \{g_{it}^\dagger\}_{i=1}^{I}\, \tilde{d}_t^\dagger,\, \tilde{t}_t^\dagger,\, \tau_t^\dagger\}_{t=0}^{\infty},\, \{\kappa_i^\dagger\}_{i=1}^{I},\, \{\kappa_j^{w\dagger},\, \kappa_j^{c\dagger}\}_{j=1}^{J},\, \{\{\tilde{w}_{jt}^\dagger\}_{j=1}^{J}, \{p_{it}^\dagger\}_{i=1}^{I},\, r_t^\dagger\}_{t=0}^{\infty}$$

and functions $d_c^\dagger(p_{it}, p_t, c_t)$, $d_w^\dagger(\tilde{w}_{jt}, \tilde{w}_t, l_t)$, $d_g^\dagger(p_{it}, p_t, g_t)$. Using the remark made above we can concentrate only on the component of the variables that responds to fiscal shocks (we deal with the technological shock analogously) and on Φ_G. Since we can determine the values of all the aggregate variables, we can also find a solution to the firm's and household's problem under perfect information. Denote the profit-maximizing price of the firm $p_{it}^{G\star}$, the optimal wage rate of the household $\tilde{w}_{jt}^{G\star}$ and the optimal composite consumption $c_{jt}^{G\star}$. We will show that the triple $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$ is a fixed point of Φ_G. Carrying out step 1 of the definition of the mapping Φ_G we obtain a triple $(p_{it}^{G\dagger},\, \tilde{w}_{jt}^{G\dagger},\, c_{jt}^{G\dagger})$ which by the definition of equilibrium is the component of $(p_{it}^{\dagger},\, \tilde{w}_{jt}^{\dagger},\, c_{jt}^{\dagger})$ that responds to fiscal shocks. Hence we arrive at the original (equilibrium) aggregate dynamics of the economy in step 2. Thus step 3 gives us back the triple $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$. Step 4 implies that $\Phi_G(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star}) = (p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$, so that $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$ is a fixed point of Φ_G.

[11] Symmetry of firms and households and the uniqueness of solution to the agents' problems assumed above imply that the decisions will be the same for all firms and the same for all households.

" $\Longleftarrow$" Suppose we have a fixed point of the mapping Φ. We will again concentrate on the mapping Φ_G and the part of the equilibrium that is associated with the fiscal shock (the proof for the technological shock is analogous). We have $\Phi_G(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star}) = (p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$. We can carry out step 1 of the definition of Φ_G to obtain $(p_{it}^{G},\, \tilde{w}_{jt}^{G},\, c_{jt}^{G})$ which we denote $(p_{it}^{G\dagger},\, \tilde{w}_{jt}^{G\dagger},\, c_{jt}^{G\dagger})$. Carrying out the optimization in step 1 we get $\{\kappa_i^{\dagger}\}_{i=1}^{I}, \{\kappa_j^{w\dagger},\, \kappa_j^{c\dagger}\}_{j=1}^{J}$ as a part of the solution. Using the analytical result proven in the Technical Appendix we get the functions $d_c^{\dagger}(p_{it}, p_t, c_t)$, $d_w^{\dagger}(\tilde{w}_{jt}, \tilde{w}_t, l_t)$ compatible with the optimization performed in step 1 and equations 2.22 - 2.28 adjusted by modifications described in 2.2.3. Using the triple $(p_{it}^{G\dagger},\, \tilde{w}_{jt}^{G\dagger},\, c_{jt}^{G\dagger})$ and equations 2.13 - 2.21 we can determine the aggregate dynamics of the economy. In this way we have defined the variables $\{\{c_{ijt}^{\dagger}\}_{i,j=1}^{I,J}, \{l_{ijt}^{\dagger}\}_{i,j=1}^{I,J}, \{\tilde{b}_{jt}^{\dagger}\}_{j=1}^{J}, \{g_{it}^{\dagger}\}_{i=1}^{I}\, \tilde{d}_t^{\dagger},\, \tilde{t}_t^{\dagger},\, \tau_t^{\dagger}\}_{t=0}^{\infty}$, $\{\kappa_i^{\dagger}\}_{i=1}^{I}, \{\kappa_j^{w\dagger},\, \kappa_j^{c\dagger}\}_{j=1}^{J}, \{\{\tilde{w}_{jt}^{\dagger}\}_{j=1}^{J}, \{p_{it}^{\dagger}\}_{i=1}^{I}, r_t^{\dagger}\}_{t=0}^{\infty}$ and functions $d_c^{\dagger}(p_{it}, p_t, c_t)$, $d_w^{\dagger}(\tilde{w}_{jt}, \tilde{w}_t, l_t)$, $d_g^{\dagger}(p_{it}, p_t, g_t)$ (denote this allocation by ($\dagger$)) We will show that ($\dagger$) is an equilibrium of the model with rational inattention. Conditions 1 and 2 of the definition of equilibrium are satisfied, because the triple $(p_{it}^{G\dagger},\, \tilde{w}_{jt}^{G\dagger},\, c_{jt}^{G\dagger})$ solves the problem of the firm and household when the optimal choice (under perfect information) is given by $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$. We have to verify that $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$ really is the optimal choice under perfect information when aggregate dynamics is given by ($\dagger$). This follows from the fact that $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$ is a fixed point of Φ_G (in particular it is the result of step 3 in the definition of Φ_G). Conditions 3, 4 and 5 of the equilibrium are satisfied trivially given the way we have defined ($\dagger$). $\square$

The Proposition and the definition of the mapping Φ provide a constructive method to search for the equilibrium numerically. We first make a guess concerning the triple $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$. We use a lag polynomial of high but finite order to represent those variables on a computer. We compute the value of Φ_G at $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$ performing the optimization of the firm, wage-setter and consumer numerically[12]. If the obtained value is close enough to $(p_{it}^{G\star},\, \tilde{w}_{jt}^{G\star},\, c_{jt}^{G\star})$, then we have found an approximation of the equilibrium. Otherwise we iterate using a linear combination of the last value and the old guess as a new guess in the next iteration until we reach the demanded precision.

2.3 Calibration

We use US quarterly data to calibrate the model. The values of a few parameters in the model will be chosen on the basis of literature and some parameters will be estimated.

Since the output elasticity of labor, α, can be interpreted as labor income share in GDP, we set $\alpha = 2/3$. The subjective discount factor, β, is related to other variables in the model by the following formula: $\frac{R}{\Pi} = \frac{1}{\beta}$. Hence, given the average nominal federal funds rate and average gross inflation rate (1965-2011), we can get $\beta = 0.988$.

Government parameters in the model include ω_B, θ_g, τ^w, θ_τ, ω plus the parameters governing the stochastic process for government spending σ_G and ρ_G.

We set $\omega_B = 6.15$, because this will imply that the debt-to-GDP ratio in the model will be 4, which is consistent with the fact that US debt is almost as high as its yearly GDP. We interpret θ_g as the share of government consumption expenditures in GDP. Hence $\theta_g = 0.2$. We use Census Bureau data to determine θ_τ, the share of taxes levied on labor (including social security) in total tax receipts (we include both federal and state and local governments). This share is approximately 2/3. We take the value of marginal personal income tax rate (for average wage workers), 0.217, (as reported by OECD) as the steady state tax rate τ^w. The above values imply that in the steady state the relation of government receipts to government spending equals 0.96 which means that the deficit-to-GDP ratio is approximately 1%. We use a simple regression model to determine ω. As the output gap we take the deviation of actual GDP from its trend obtained by applying a HP filter. We get $\omega = 2.58$. It means that fiscal policy is anticyclical: an increase in GDP of 1% increases the relation of government receipts to government spending by 2.58 percentage points or, given the steady state budget deficit, decreases the deficit-to-GDP ratio from 1 to 0.4%. To determine the parameters of government spending shock

[12] To solve those problems numerically in finite time we have to make some concessions, for example instead of using the MA representation based on the lag polynomial with a big number of coefficients, we use the ARMA(p,q) representation where p and q are small. We also compute the information flow 2.33 for large but finite T.

we detrend the time series of logarithm of government spending and estimate a simple regression which yields $\rho_G = 0.97$ and $\sigma_G = 0.02$[13].

To determine the values of $\tilde{\theta}$ and $\tilde{\eta}$, the elasticities of substitution between different types of goods and labor, respectively, we use data on average mark-ups[14]. We set $\tilde{\theta} = 4$, $\tilde{\eta} = 4$ which is consistent with the mark-up of 33%.

It is difficult to find a sensible value for the intertemporal elasticity of substitution, inverse of γ, and for the labor elasticity, inverse of ψ. For example, in a simple DSGE model only small values of γ yield a positive response of labor supply to technological shocks which is consistent with empirical evidence. However, micro data suggests that the elasticity is much smaller than one, implying a big value for γ. We set $\gamma = 2$ to match the estimates of intertemporal elasticity of substitution in Basu and Kimball (2002) which are close to 0.5 (depending on the model specification). Following Mackowiak and Wiederholt (2010), we set $\psi = 0$ (utility is linear in labor)[15].

We use a standard Solow decomposition method to obtain the estimates of the level of technology and a simple regression to determine the parameters governing the technological process. Since there is no capital in the model, we only use effective labor (employment times hours worked) as a factor of production. This procedure yields the standard deviation of the shock of $\sigma_A = 0.0069$ and the autocorrelation coefficient of $\rho_A = 0.951$. The parameters of the idiosyncratic technological shock are taken from Mackowiak and Wiederholt (2010). Using price micro data they obtain $\rho_I = 0.3$ and $\sigma_I = 0.18$.

We have to determine the coefficients in the monetary policy equation 2.21. It turns out that $\phi_\pi > 1$ and $\phi_y > 0$ are necessary conditions for uniqueness of the equilibrium. We take $\phi_\pi = 1.07$, $\phi_y = 0.11$, $\rho_R = 0.92$ using the estimation results obtained by Giammarioli *et al.* (2008). We choose these particular parameter values to model monetary policy as relatively mild. This implies that fiscal policy effects are more distinct.

Finally, we must find values for μ and λ, the marginal cost of information for the firm and for the household. This is tricky as there are no direct equivalents of these parameters in economic data, especially that we want to interpret cost of information as welfare cost rather than monetary cost. We thus propose the following procedure. We will exploit the fact that parameters μ and λ effect directly the degree of price and wage stickiness in equilibrium. We treat the technological shock as a benchmark shock. We solve a model which is identical to the model with perfect information except for the existence of nominal stickiness in the form of Calvo price and wage setting. We set the Calvo parameters, ζ_F - the probability that the firm cannot change the price and $\zeta_H-$the probability that the household cannot change the wage rate, to match a consensus in the New Keynesian DSGE literature. It seems that $\zeta_F = \zeta_H = 0.6$ is near that consensus. We then choose the parameters μ and λ numerically to minimize the distance between impulse response functions of price and wage reaction to a technological shock in the two models (one with perfect information and Calvo-style stickiness and one with rational inattention). One advantage of this approach is that the stickiness generated by rational inattention in the case of the technological shock will be consistent with empirical evidence.

3 Results

In this section we discuss our main results. We focus on the fiscal shock as the effects of idiosyncratic and aggregate technological shocks where analyzed in much detail in Mackowiak and Wiederholt (2010) and Mackowiak and Wiederholt (2007). Nevertheless we present some graphs (impulse response functions to an aggregate technological shock) in the Appendix. Our findings confirm the conclusions of Mackowiak and Wiederholt (2010). Under rational inattention the prices are sticky, although apart

[13]In fact the estimated parameters vary substantially depending on the time period we choose. We take, somewhat arbitrarily, an average value of those results. We later check the implications of changing the variance of the fiscal shock.

[14]See for example: Gali *et al.* (2011), Allayannis and Ihrig (2001)

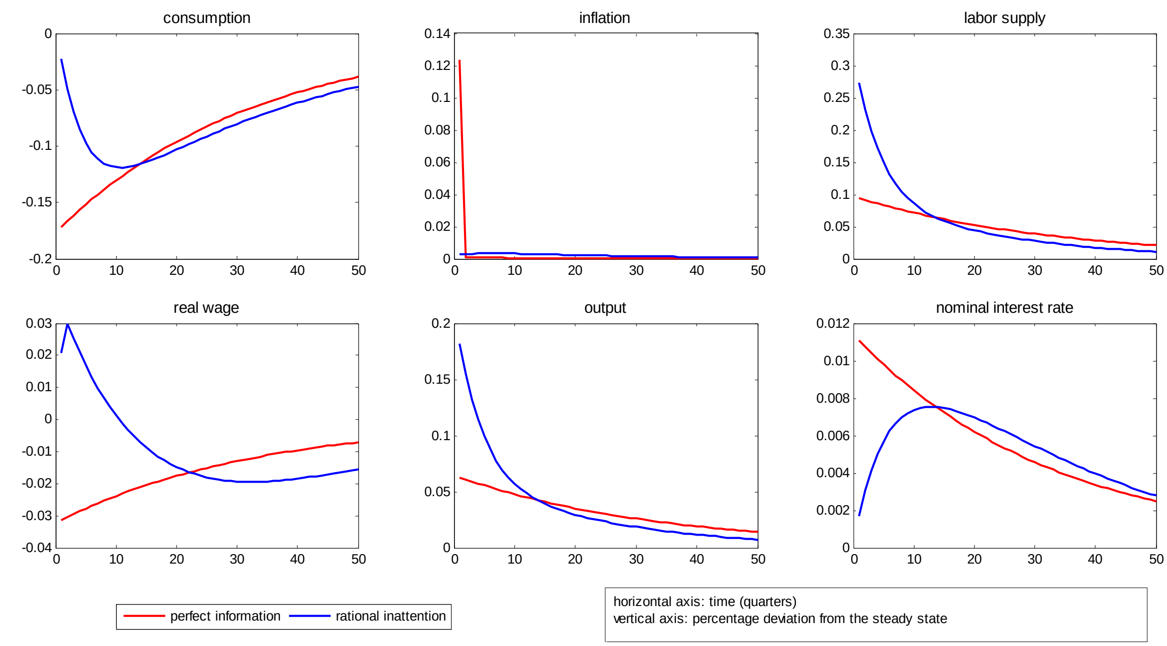
[15]It is important to stress, that this assumption is not necessary to solve our model, unlike in the case of the solution method proposed by Mackowiak and Wiederholt (2010).

from information processing constraints, there are no frictions in the economy[16]. The reaction of key economic variables (like consumption or output) is humped-shaped. Also, firms adjust prices much quicker in the cace of idiosyncratic technological shocks than in the case of aggregate technological shocks - a pattern observed in microeconomic data[17].

3.1 Impulse response to a fiscal shock

Figure 3.1 presents the reaction of main economic variables to a 1% deviation of government spending from its steady state under basic calibration. The results are contrasted with the equilibrium under perfect information to better illustrate the mechanics of rational inattention.

Figure 3.1: Response to 1% shock to government spending



Consider first the reaction of consumption. When all agents have perfect information, consumption falls immediately on shock and then returns monotonically to its steady state. This is caused by the negative wealth effect of increased government spending. Under rational inattention the wealth effect is still in place, but since households allocate only limited attention to observing the fiscal shock, they adjust their consumption expenditures gradually. Thus the reaction of consumption is hump-shaped and initially weaker than under perfect information. After about 12 quarters consumption under rational inattention falls short of consumption under perfect information. This is caused by the fact that the households increase their debt (relative to the debt they would have under perfect information) in the initial periods when consumption is higher and since suboptimal bond holding has the tendency to grow over time (see equation 2.49), they eventually limit their consumption to repay that debt. Summing up, rational inattention reduces the crowding-out effect in the short and middle term.

[16] In particular price and wage stickiness does not depend on an arbitrary mechanism like the Calvo-style pricing and wage-setting which rather assume than imply stickiness

[17] In fact the response to idiosyncratic technological shock under our calibration is visually the same as under perfect information

Consider next the reaction of firms. Under perfect information firms increase their prices immediately after the shock. When firms are rationally inattentive, the prices hardly react to a fiscal shock. This can be explained by the fact that a mistake made by the firm when setting the price after an increase in government spending is cheap in terms of expected loss in profits. Indeed, government spending only accounts for a relatively small fraction of aggregate demand and the partial crowding-out of consumption further reduces the effect on aggregate demand. Note that this irresponsiveness of prices to fiscal shock coexists with a much stronger reaction of prices to technological shocks (especially idiosyncratic). This is a property of the model that cannot be reproduced by conventional Calvo-style pricing.

To explain the reaction of wages and labor note first that under perfect information households set lower wages to induce firms to demand more labor and make up for the decrease in wealth. This is despite the fact that increased government spending entails higher income taxes. As we show in the Appendix, the optimal wage rate depends positively on consumption and aggregate labor supply. Under rational inattention consumption is higher as a consequence of information processing constraints of households, but there is an additional reinforcement effect. Higher consumption means higher demand for the goods produced by firms, which is also boosted by lower prices set by firms. As a result firms increase production significantly and demand more labor. This leads to an increase in wages.

In the absence of technology shocks output rises in accordance with labor supply. The resulting fiscal multiplier under rational inattention is much higher than under perfect information. The contemporaneous multiplier is 0.91 compared to 0.31 when agents have perfect information. This value is very close to estimates obtained by Blanchard and Perotti (2002) and other authors. We also conduct the following experiment: we introduce Calvo price and wage setting to the model with perfect information and search for the value of parameters ζ_F and ζ_H that maximize the fiscal multiplier. We find out that under our simple specification and calibration Calvo stickiness cannot produce a multiplier larger than 0.51[18]. We can conclude that rational inattention is a simple and elegant explanation of why the multiplier can be close to one. We show in the next section that it can also account for other values of the multiplier.

Figure 3.2 gives some insight into how rational inattention affects individual decisions of households and firms. We plot the actual decision under information processing constraints against the optimal decision (under perfect information) but when aggregate dynamics is given by the equilibrium of the model with rational inattention[19]. Note the difference between Figure 3.1 and Figure 3.2. In the first case the red lines indicate the equilibrium allocation that arises when all agents have perfect information. The red lines in Figure 3.2 show what an agent would choose if he had perfect information and all other agents acted under rational inattention. By comparing these two graphs we learn how interactions between agents move the equilibrium under rational attention away form the perfect information case.

Notice for example that the optimal price change for a firm in the rational inattention economy is much smaller than for a firm in the perfect information environment. This is due to a complementarity in price setting. When other firms issue lower prices, the firm, interested only in relative prices, also sets a lower price. The optimal wage rate in Figure 3.2 illustrates the point made earlier that higher consumption and lower prices increase employment and make a rise in wages optimal. Under information processing constraints the wage rises only slightly resulting in the behavior of wages depicted in Figure 3.1. Finally, the optimal reaction of consumption is somewhat weaker than in the case when all agents have perfect information. This is a consequence of a different behavior of the interest rate. In the rational inattention equilibrium, the interest rate is lower due to lower inflation. This induces households to substitute future consumption for current consumption. As a result, optimal consumption is higher in the initial periods.

[18]One can increase the multiplier easily by increasing γ, the inverse of the intertemporal elasticity of substitution. But a value of γ much higher than 2 would be at odds with empirical evidence.

[19]Whenever we make a simulation of this kind, we assume that the realization of noise is zero in every period.

Figure 3.2: Optimal reaction under perfect information and under rational inattention

Note one important limitation of rational inattention in the context of fiscal policy. Many empirical studies based on the VAR approach find that consumption reacts positively to fiscal shocks (see for example Mountford and Uhlig, 2008, Blanchard and Perotti, 2002, Fatas and Mihov, 2001). Without further modifications our model cannot explain why consumption would rise in response to a fiscal shock. This is clear when we combine two facts. First, under perfect information it is optimal for the households to decrease consumption. Hence, if rational inattention were to account for rising consumption, the deviation from the optimal choice would have to be large and hence the information flow would have to be small. The second fact is connected with the properties of mutual information. Information flow is small when the variance of noise is big in comparison to the variance of the pure signal. What matters for information flow is therefore the magnitude rather than the sign of the coefficients in the lag polynomial. A positive reaction of consumption cannot be therefore optimal, because decreasing the positive coefficients in the lag polynomial towards zero would both decrease the expected utility loss and the information flow.

3.2 The effect of changing the variance of the shock

In this section we analyze the equilibrium of the model with rational inattention for different values of the variance of the fiscal shock. The variance of the shock plays a crucial role in determining the optimal allocation of attention. Since the objective of the firm and the household is a quadratic function of the difference between optimal and actual decision, and since those decisions are expressed as MA processes based on innovations to government spending ε_t^G, the variance of ε_t^G enters directly into 2.35, 2.42 and 2.47. When the variance increases, the potential welfare or profit losses are greater and the agents allocate more attention to observing the fiscal shock. This has implications for aggregate dynamics.

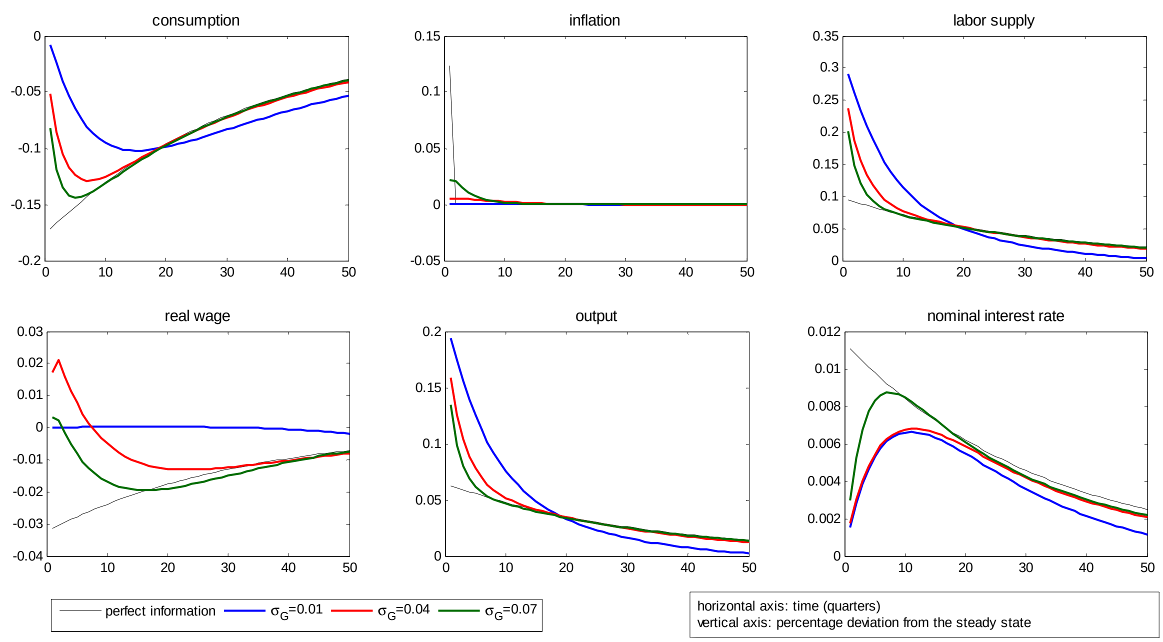
Figure 3.3 presents the reaction of the economy to a 1% deviation of government spending from its steady state for different values of standard deviation of the fiscal shock[20] ($\rho_G = 0.01$, 0.04 or 0.07). When the variance of the shock is low ($\sigma_G = 0.01$) both firms and households pay very little attention to fiscal shocks. The equilibrium attention level κ is approximately 0 for the price-setting decision of

[20]This may be somewhat misleading. The actual shock in all three cases is normalized to 1%, the difference lies in the characterization of the stochastic process governing the behavior of government spending.

firms, 0.0006 for wage-setters and 0.13 for composite consumption decision of consumers[21]. As a result consumption reacts slowly and wages and prices are practically completely stiff. This implies that the crowding-out effect is significantly reduced and the rise in output is greater. When the variance of the shock rises, so does the equilibrium attention capacity. For firms κ equals 0.1 when $\sigma_G = 0.04$ and 0.18 when $\sigma_G = 0.07$. The corresponding numbers for the wage-setter are 0.14, 0.18 and 0.5, 0.85 for the consumer. Consumption reacts more quickly and strongly and the prices exhibit a slight increase. Wages fall eventually, although they still have a tendency to rise right after the shock. The crowding-out effect gets stronger.

The contemporaneous fiscal multiplier is 0.97 when $\sigma_G = 0.01$, 0.79 when $\sigma_G = 0.04$ and 0.67 when $\sigma_G = 0.07$. The effectiveness of fiscal policy (its ability to influence real output) depends on the variance of the fiscal shock. Note that the parameter σ_G^2 that the agents use in their optimization should rather be interpreted as expected or perceived variance. Therefore, the perceived stability of fiscal policy may influence its effectiveness. For example, when fiscal policy is seen as stable, agents pay little attention to fiscal shocks and the crowding-out effects are small. However, when fiscal policy gets volatile, agents allocate more attention to observing fiscal shocks and its effectiveness is reduced. This is an important conclusion in times like the financial crisis when we witnessed massive fiscal expansions. The theory of rational inattention suggests that they might have been less effective, because agents allocated more attention to making decisions related to fiscal policy.

Figure 3.3: Response to 1% shock to government spending (different variance of the shock)



There are some formal problems with this reasoning. First, one should be more precise about the interpretation of σ_G^2 and try to explain how the agents formulate their expectations about the actual variance of the fiscal shock. Second, the form of households' and firms' problems does not allow reallocation of attention. When we compared the three cases above, we simply treated them as three different scenarios. It would be interesting to see how the agents behave when the volatility of fiscal policy changes in the course of economic life. Third, in order to account for different values of the

[21] Note by the way that these numbers suggest that the choice of composite consumption is by far more important for the household than the choice of wage rate.

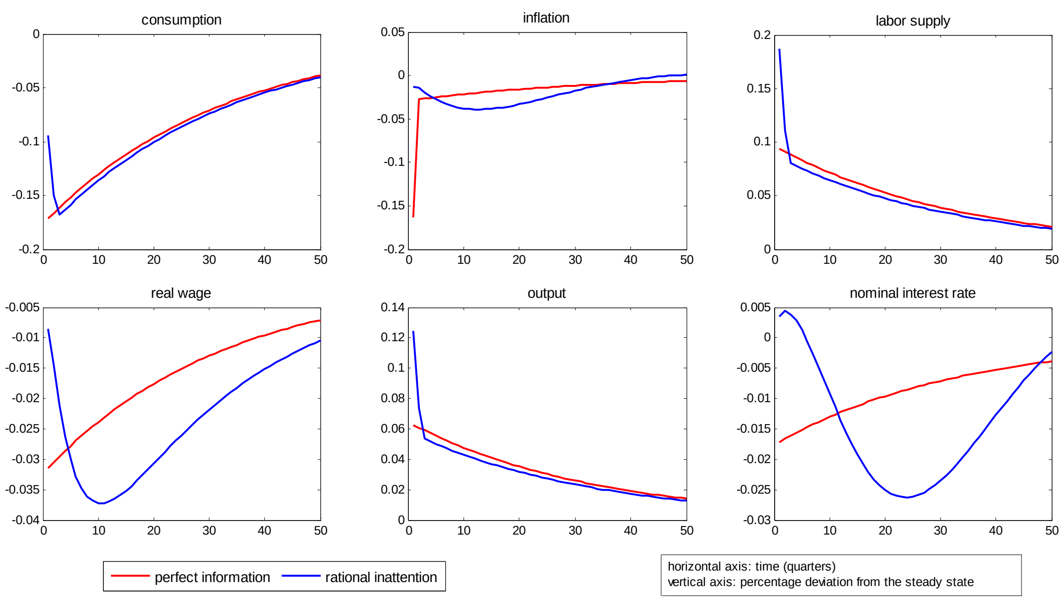
fiscal multiplier, the variance of the shock had to be changed by quite a large amount. To explain the values of the multiplier closer to 0.5 the variance would have to be unreasonably high. We address all those issues in Section 4.

Note that we could conduct a very similar analysis changing the cost of information flow, λ and μ. Lower cost of processing information induces agents to increase the level of attention, thus reducing the effects of an increase in government spending. Since parametrs λ and μ are tricky to estimate and have no good equivalents in the real economy, we cannot trace the way they evolved in the last decades, but with the fast pace of technological progress that we experiencd (widespreading computer and Internet usage in particular) we could probably agree, that they tend to fall. We come to a surprising conclusion. The effectiveness of fiscal policy may decrease with the development of efficient methods to process information, or equivalently, over time. Such a phenomen was actually described in the literature. Bilbiie *et al.* (2006) show that the effects of fiscal expansions are weaker after 1980 than they used to be earlier.

3.3 Fiscal-monetary policy interactions

We now change the parameters of monetary policy rule 2.21 and analyze the consequences for equilibrium under rational inattention. This exercise casts some light on the interactions between fiscal and monetary policy under information processing constraints.

Figure 3.4: Response to 1% shock to government spending (alternative monetary policy rule)



We set $\phi_\pi = 1.5$ and $\phi_y = 0.5$. Consider Figure 3.4. Under perfect information, prices and the nominal interest rate fall as a response to shock. Other variables behave almost identically as in the case of baseline calibration (see Figure 3.1). This is due to the fact that the real interest rate that matters for consumption is similar in both settings. However, the allocation under rational inattention is significantly different.

When monetary policy is more aggressive, prices fall as a response to a fiscal shock. Since firms are only interested in relative prices, it is optimal for an individual firm to lower its price as well. As a consequence of information processing constraints prices (at least initially) fall less than they should. This in turn implies that the nominal (and real) interest rate is higher and this dampens consumption.

Hence, when monetary policy is restrictive, rational inattention on the side of firms increases the crowding-out effect. Since rational inattention on the side of households decreases crowding-out, we have two effects working in opposite directions. As a result, the fiscal multiplier is reduced to 0.7 (compared to 0.91). Consumption, output and labor supply fall short of their perfect information equivalents much sooner.

Furthermore, more aggressive monetary policy alters the allocation of attention. Under baseline calibration, equilibrium attention level κ is 0.07 for the price-setting decision of firms, 0.1 for wage-setters and 0.24 for composite consumption decision of consumers. When we change the parameters of monetary policy the numbers become 0.29, 0.25 and 0.39. That explains why the reaction of wages in Figure 3.4 is much closer to the perfect information case than it is in Figure 3.1. This also means that under more restrictive monetary policy pricing errors made by firms become much more costly in terms of expected profits.

We can conclude that under rational inattention, aggressive monetary policy (in this context aggressive means inducing the prices to fall after fiscal shock) has a stronger adverse influence on the effectiveness of fiscal policy.

4 Dynamic allocation of attention

This section introduces a new approach to solving the attention allocation problem of agents. To the best of our knowledge, nobody has yet studied the possibility that agents may allocate their attention in every period. It has rather been assumed (as in this paper) that agents choose their level of attention in the initial period once and for all. It turns out that allowing for dynamic allocation of attention increases the potential of the model to explain various economic phenomena. For example, Mackowiak and Wiederholt (2007) conclude that rational inattention cannot explain why prices sometimes stay fixed for some periods after a shock before they react (as indicated by micro data). We show that when agents choose attention level dynamically, prices may stay practically fixed for some time and then react to a shock quite strongly. Besides, the modifications proposed here address the concerns aired in section 3.2 and formalize the idea that the size of the shock matters.

An additional motivation for this section is a simple observation, that economic conditions are likely to influence the amount of time and energy that agents sacrifice in order to make economic decisions. When the economy becomes volatile, as it did for example during the financial crisis, the optimal allocation of attention may change, simply because the expected welfare or profit losses caused by bad decisions become greater. On the other hand, when the economy is stable for a long time (say, there are no fiscal shocks), the level of attention may decrease over time. Our modification captures this intuition.

The model presented below is a little bit stylized, but the aim of this section is to point out a problem in an approachable way. We leave a rigorous treatment and incorporating this idea into general equilibrium for future research[22].

Let us start with a straightforward observation. Assume that there is a shock $\varepsilon_0 \sim N(0, \sigma^2)$ in period 0 and that afterwards agents observe a signal of the form $b_0\varepsilon_0 + c_0\nu_0$, where $\nu_0 \sim N(0, 1)$ represents noise. This is a situation that takes place in the model presented above one period after the agents have chosen their allocation of attention[23] (actual decision plays the role of a signal). Notice that in period 1, we have

$$\mathbb{E}[\varepsilon_0^2 | b_0\varepsilon_0 + c_0\nu_0] = \sigma^2 \left[\frac{c_0^2}{b_0^2\sigma^2 + c_0^2} + \frac{b_0^2\sigma^2}{b_0^2\sigma^2 + c_0^2} \frac{(b_0\varepsilon_0 + c_0\nu_0)^2}{b_0^2\sigma^2 + c_0^2} \right] \tag{4.1}$$

Depending on the signal-to-noise ratio and the realization of ε_0 and ν_0 the above number can be greater or smaller then σ^2 which is the expectation of ε_0^2 before observing the signal (in period 0, when allocation of attention is chosen). Note that the expressions $\frac{c_0^2}{b_0^2\sigma^2+c_0^2}$ and $\frac{b_0^2\sigma^2}{b_0^2\sigma^2+c_0^2}$ play the role of

[22]The reason we don't do that now is that using the same method to search for equilibrium as before but with dynamic allocation of attention would yield an extremely time-consuming algorithm that would be of little use in practice.

[23]Given that all the realizations of the shock before period 0 were 0

weights. For example when $c_0 = 0$, $\mathbb{E}[\varepsilon_0^2|b_0\varepsilon_0 + c_0\nu_0] = \varepsilon_0^2$. This shows that when the realization of the shock in a given period is high and optimal decision in the current period depends on the realizations of shocks in previous periods, agents would like to reallocate attention, simply because they realize that they underestimated the magnitude of potential losses due to suboptimal decision.

We will make an assumption concerning the way agents estimate and forecast the variance of the shock σ^2. This problem was hinted at in section 3.2. So far we have implicitly assumed that agents know the actual, objective variance σ^2. Instead, we now assume that agents have a prior belief about variance and let the agents formulate expectations of future ε_t^2 on the basis of already observed signals. By signals we mean actual decisions of the form 2.39, 2.45 or 2.51 which provide information about past realizations of ε_t. Let y_t denote all signals observed between periods 0 and t (including t) (for example for a wage setter we have $y_t = (\tilde{w}_{j0}^{\dagger}, \tilde{w}_{j1}^{\dagger}, ..., \tilde{w}_{jt}^{\dagger}))$. To keep the model simple we assume that agents received n signals suggesting that the variance is σ^2 before period 0 and in every period t they set

$$\mathbb{E}[\varepsilon_{t+s}^2|y_t] = \frac{1}{n+t}\left(n\sigma^2 + \sum_{i=0}^{t-1}\mathbb{E}[\varepsilon_i^2|y_t]\right) \quad s = 0, 1, 2, ... \tag{4.2}$$

This means that agents assume that future variance of the shock will be equal to the variance estimated from available data. Parameter n can be seen as the persistence of the prior belief about the variance. The estimation technique is primitive and we do not claim that it is optimal. However, in our view it illustrates the problem quite well.

We can now state the optimization problem. Take for example the problem of the firm 2.35 - 2.41. We simply assume that the firm solves this problem in every period $t \geq 0$, with the expectation operator changing to $E_{i,t}$ accordingly. The firm computes $E_{i,t}\varepsilon_s^2$ using conditioning analogous to 4.1 (taking into account all available signals) for $s \leq t$ and expectation technology 4.2 for $s > t$. Note that in every period the firm chooses whole lag polynomials (see equation 2.39), but only the first coefficients matter for behavior, as new lag polynomials will be chosen in the next period. We can modify the problems of the wage-setter and of the consumer in a fully analogous way.

There are several shortcomings of both theoretical and practical nature of the proposed approach. For example, the mutual information operator of the form 2.33 measures average per-period information flow between two stochastic processes. The actual information flow in a given period may be higher or lower. When we solve the attention problem dynamically we continue to use the operator 2.33, although the chosen stochastic processes only provide information flow in the current period (in the next period new stochastic processes are chosen). This means that the actual information flow may not satisfy the constraint 2.41. One way to fend off such criticism would to be argue that agents solving the attention problem in any given period t do not expect to change their decision in the following periods and treat the chosen stochastic processes as describing their decision from now on. It is the signals in consecutive periods that make them change their mind as they realize that their estimation of the real variance may have been wrong. This also means that agents do not behave strategically taking into account the influence of the today's choices on future accuracy of signals. Another problem arises due to the fact that we always have to make an assumption concerning the realizations of shocks before the initial period in which the decision is made. So far we have assumed that the realization was a vector of zeros. Now this issue becomes more cumbersome. Not only is this assumption contradictory to the fact that agents observe n signals suggesting that the variance is σ^2 before period 0, but it also causes the attention level to rise over time. To see this note that in every period $t \geq 0$ additional terms enter into the objective 2.35 as the non-trivial (non-zero) history of the shocks $(\varepsilon_0, \varepsilon_1, ..., \varepsilon_{t-1})$ gets longer. Because the information flow constraint does not depend on the particular realizations of shocks, the expected loss in profits due to suboptimal decisions rises in comparison to the information processing costs. This induces firms to increase the level of attention. To get rid of this effect, we make an (admittedly schizophrenic) assumption that in every period $t \geq 0$ all terms $E_t\varepsilon_s^2$ for $s < t$ are zero in the objective 2.35[24]. Difficult as this assumption may be to justify, it makes the presentation of the results far more clear.

[24] Although the firm uses the actual value of $E_t\varepsilon_s^2$ to estimate the variance of the shock. That's why we called the assumption schizophrenic.

The remainder of this section is devoted to discussing two applications of the proposed modification. We first study dynamic allocation of attention by a consumer choosing composite consumption. We then check how it changes the price-setting behavior of firms.

4.1 Dynamic allocation of attention by the household

We study the optimization problem of the consumer (choosing composite consumption) when the aggregate conditions are given by the benchmark rational inattention equilibrium described in section 3.1. We conduct an experiment analogous to the one depicted in Figure 3.2, but we now let the household choose attention level κ dynamically as described in the previous section. Two cases are considered. In the first scenario the fiscal shock is very big, $\varepsilon_1^G = 0.3$ In the second, we consider a small shock, $\varepsilon_1^G = 0.001$. The results are presented in Figures 4.1 and 4.2[25].

Figure 4.1: Dynamic allocation of attention - big shock

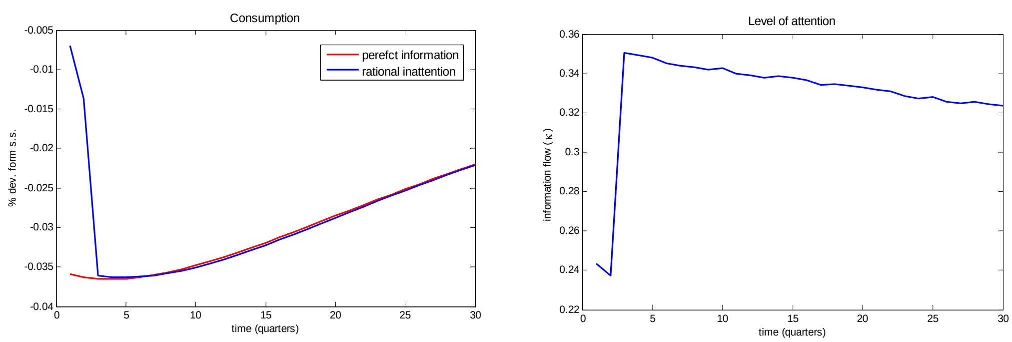
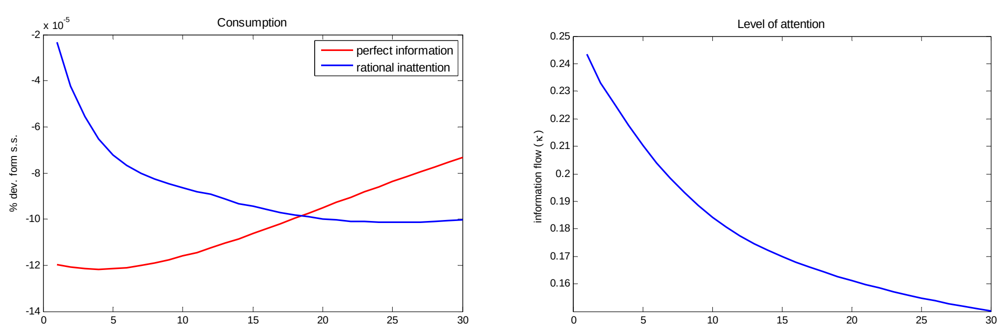


Figure 4.2: Dynamic allocation of attention - small shock



[25]We set $n = 5$.

Consider first the reaction of consumption when the fiscal shock is big. Consumption returns to its optimal level very quickly, already in the third period. This is caused by the increase in attention level κ from the initial 0.24 to 0.35. Note that attention capacity falls in the second period although the shock happens in period 1. An explanation of this surprising fact is quite simple. The agent uses her decisions as signals. In the first period, the reaction of consumption is very weak and it thus provides little information about ε_1^G. More formally, the first period decision has a small signal-to-noise ratio and thus the consumer puts little weight on the observation made in the first period. In the second period, consumption reacts more strongly to the first period shock (providing a more accurate signal) and the agent is able to estimate the value of ε_1^G more precisely. As he realizes that the square of ε_1^G was much greater than his initial belief σ_G^2 he updates his estimate of future variance according to 4.2. This induces him to increase attention level. Because the realization of shock is zero in all periods $t > 1$, the attention capacity then gradually decreases.

When the shock is small, the picture is completely different. See Figure 4.2. Now the agent systematically lowers her expectation of the variance of the shock as new signals arrive. The attention level falls over time. As a result, consumption takes much more time to reach the level it would reach if the household had perfect information.

This suggests that the effectiveness of fiscal policy may depend on the size of the shock. As long as the shock is small, agents allocate little attention to observing it, consumption stays high for many periods and the effect on real output is greater. However when fiscal authorities want to boost the economy tapping a huge fiscal stimulus or increase spending significantly due to war expenditures, agents increase their attention capacity and consumption falls very quickly weakening the impact of fiscal expansion.

We thereby proposed an explanation of why different empirical studies find varied values of the fiscal multiplier. When the focus of the study is on war episodes and defense spending, the shocks in question are bigger (we find that the variance of log-deviated defense spending is larger than the variance of non-defense spending). The estimates should therefore be lower. This does seem to be the case in the literature, as indicated in the Introduction.

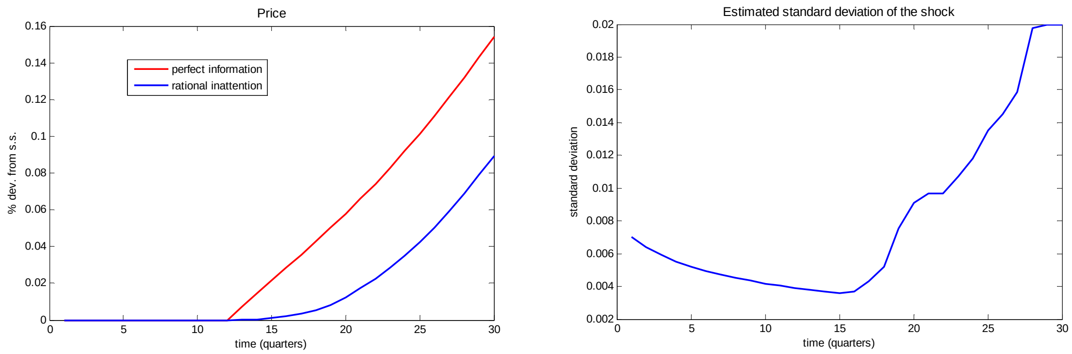
4.2 Dynamic allocation of attention by the firm

We now examine a behavior of the firm choosing its price when aggregate conditions are given by the benchmark rational inattention equilibrium described in section 3.1. In order to show that prices may stay practically constant for some time and then react we consider the following situation. First, there are no shocks for a couple of periods. Then, for some reason, the economy gets volatile and we observe a series of shocks of considerable magnitude. For the sake of legibility of the graph we assume that the realization of shock in every consecutive period is equal and positive ($\varepsilon_t^G = 0.2$), but this is not essential for the main conclusion to hold (what is important is that the shocks are large in absolute value). Figure 4.3 presents the results of this experiment. Note that this time, unlike in Figure 3.2, we plot the price level, not the price change. We also depict the estimated standard deviation of the shock instead of attention level. Obviously, these two quantites are closely related.

The price level reacts to the shocks with a lag of about three periods. This is caused by the fact that the attention level is very low after a long time of lack of movement in the economy when the first shock strikes. Only gradually does the firm realize that something is going on and it increases its attention capacity. The process is slow at first because a low level of attention means that the decisions of the firm convey very little information about the state of the economy. Then it is sped up by the interaction of higher estimated standard deviation and higher level of attention. Higher expected variance induces the firm to increase attention capacity and higher attention capacity increases the informational content of the decision of the firm and triggers faster variance updates. As a result, prices rise.

Although the above example is very stylized, we believe that it illustrates quite well the reasons for which firms may not change prices. When prices stay fixed for some time because there is no need

Figure 4.3: Dynamic allocation of attention by the firm



to change them (there are no shocks), decision makers may get used to this situation and start to pay less attention to market conditions (lower their attention capacity). When a shock finally occurs, it thus takes them some time to notice it. Hence prices react with a lag.

5 Conclusions

We presented a dynamic stochastic general equilibrium model with rational inattention and studied the implications of information processing constraints for fiscal policy. We showed that rational inattention decreases the crowding-out effects of fiscal expansions and can provide an explanation of why the value of the fiscal multiplier may change over time and why its estimates based on war episodes and defense spending are generally lower that others. It also casts new light on fiscal-monetary policy interactions. An innovative approach to look at attention allocation has been proposed - we argue that allowing the agents to choose the allocation of attention every period increases the potential of the model to account for various economic phenomena. One of the main conclusion can be summarized in a statement that the size of the shock matters.

The model has some drawbacks. First, dividing the household's problem into two subproblems, although not uncommon in economic literature, is somewhat unjustified in the context of rational inattention as it requires a strong assumption that processing information to decide about wages and processing information to choose consumption are separate and unrelated activities. Such assumptions are often made in models with rational inattention, but also often criticized (see Sims, 2010). Second, the model cannot explain why consumption could rise in response to an increase in government spending. There are several explanations of this phenomenon in the literature (rule-of-thumb consumers - Gali *et al.*, 2007, non-separable utility function - Monacelli and Perotti, 2008, deep habits - Jacob, 2010). It would be interesting to see how rational inattention interacts with one of these concepts. Third, there is no capital in the model. Adding capital to the model is unlikely to change the main conclusions of this paper, but there is an open question of how investment would react to a fiscal shock under rational inattention.

Extending the model is troublesome because of lack of computationally effective methods to search for equilibrium. Finding such a method is a challenge for future research. Related to that issue is the problem of incorporating the concept of dynamic allocation of attention into a general equilibrium model.

References

Allayannis, G. and Ihrig, J. (2001) Exposure and markups, *Review of Financial Studies*, **14**, 805–35.

Barro, R. J. (1981) Output effects of government purchases, *Journal of Political Economy*, **89**, 1086–1121.

Barro, R. J. and Redlick, C. J. (2009) Macroeconomic effects from government purchases and taxes, NBER Working Papers 15369, National Bureau of Economic Research, Inc.

Basu, S. and Kimball, M. (2002) Long-run labor supply and the elasticity of intertemporal substitution for consumption, unpublished Manuscript, University of Michigan.

Batchuluun, Altantsetseg, Yulei, L. and Young, E. R. (2007) Portfolio choice with information-processing limits, unpublished Manuscript, University of Virginia and Hong Kong University.

Baxter, M. and King, R. G. (1993) Fiscal policy in general equilibrium, *American Economic Review*, **83**, 315–34.

Bilbiie, F. O., Meier, A. and Muller, G. J. (2006) What accounts for the changes in u.s. fiscal policy transmission?, Working Paper Series 582, European Central Bank.

Blanchard, O. and Perotti, R. (2002) An empirical characterization of the dynamic effects of changes in government spending and taxes on output, *The Quarterly Journal of Economics*, **117**, 1329–1368.

Blanchard, O. J. and Kahn, C. M. (1980) The solution of linear difference models under rational expectations, *Econometrica*, **48**, 1305–11.

Burnside, C., Eichenbaum, M. and Fisher, J. D. (2000) Assessing the effects of fiscal shocks, NBER Working Papers 7459, National Bureau of Economic Research, Inc.

Calvo, G. A. (1983) Staggered prices in a utility-maximizing framework, *Journal of Monetary Economics*, **12**, 383–398.

Christiano, L., Eichenbaum, M. and Rebelo, S. (2009) When is the government spending multiplier large?, NBER Working Papers 15394, National Bureau of Economic Research, Inc.

Cover, T. M. and Thomas, J. A. (2006) *Elements of information theory*, 0471241954.

Edelberg, W., Eichenbaum, M. and Fisher, J. D. (1999) Understanding the effects of a shock to government purchases, *Review of Economic Dynamics*, **2**, 166–206.

Fatas, A. and Mihov, I. (2001) The effects of fiscal policy on consumption and employment: Theory and evidence, CEPR Discussion Papers 2760, C.E.P.R. Discussion Papers.

Fueki, T., Fukunaga, I. and Saito, M. (2011) Assessing the effects of fiscal policy in japan with estimated and calibrated dsge models, Tech. rep.

Gali, J., Lopez-Salido, J. D. and Valles, J. (2007) Understanding the effects of government spending on consumption, *Journal of the European Economic Association*, **5**, 227–270.

Gali, J., Smets, F. and Wouters, R. (2011) Unemployment in an estimated new keynesian model, NBER Working Papers 17084, National Bureau of Economic Research, Inc.

Giammarioli, N., Lambertini, L. and Onorante, L. (2008) Is u.s. fiscal policy optimal?, Working Papers 200802, Center for Fiscal Policy, Swiss Federal Institute of Technology Lausanne.

Hall, R. E. (2009) By how much does gdp rise if the government buys more output?, NBER Working Papers 15496, National Bureau of Economic Research, Inc.

Huang, L. and Liu, H. (2007) Rational inattention and portfolio selection, *Journal of Finance*, **62**, 1999–2040.

Jacob, P. (2010) Deep habits, nominal rigidities and the response of consumption to fiscal expansions, Working Papers of Faculty of Economics and Business Administration, Ghent University, Belgium 10/641, Ghent University, Faculty of Economics and Business Administration.

Linnemann, L. and Schabert, A. (2003) Fiscal policy in the new neoclassical synthesis, *Journal of Money, Credit and Banking*, **35**, 911–29.

Luo, Y. (2006) Rational inattention, portfolio choice, and the equity premium, Computing in Economics and Finance 2006 56, Society for Computational Economics.

Luo, Y. (2008) Consumption dynamics under information processing constraints, *Review of Economic Dynamics*, **11**, 366–385.

Luo, Y. and Young, E. R. (2009) Rational inattention and aggregate fluctuations, *The B.E. Journal of Macroeconomics*, **9**, 14.

Mackowiak, B. A. and Wiederholt, M. (2007) Optimal sticky prices under rational inattention, CEPR Discussion Papers 6243, C.E.P.R. Discussion Papers.

Mackowiak, B. A. and Wiederholt, M. (2010) Business cycle dynamics under rational inattention, CEPR Discussion Papers 7691, C.E.P.R. Discussion Papers.

Martins, G. and Sinigaglia, D. (2009) Real business cycle dynamics under rational inattention, MPRA Paper 14089, University Library of Munich, Germany.

Matejka, F. (2010) Rationally inattentive seller: Sales and discrete pricing, CERGE-EI Working Papers wp408, The Center for Economic Research and Graduate Education - Economic Institute, Prague.

Monacelli, T. and Perotti, R. (2008) Fiscal policy, wealth effects, and markups, NBER Working Papers 14584, National Bureau of Economic Research, Inc.

Mondria, J. (2010) Portfolio choice, attention allocation, and price comovement, *Journal of Economic Theory*, **145**, 1837–1864.

Moscarini, G. (2004) Limited information capacity as a source of inertia, *Journal of Economic Dynamics and Control*, **28**, 2003–2035.

Mountford, A. and Uhlig, H. (2008) What are the effects of fiscal policy shocks?, NBER Working Papers 14551, National Bureau of Economic Research, Inc.

Nieuwerburgh, S. V. and Veldkamp, L. (2008) Information acquisition and under-diversification, NBER Working Papers 13904, National Bureau of Economic Research, Inc.

Paciello, L. (2009) Monetary policy activism and price responsiveness to aggregate shocks under rational inattention, MPRA Paper 16407, University Library of Munich, Germany.

Paciello, L. (2010) The response of prices to technology and monetary policy shocks under rational inattention, Tech. rep.

Peng, L. and Xiong, W. (2005) Investor attention: Overconfidence and category learning, NBER Working Papers 11400, National Bureau of Economic Research, Inc.

Ramey, V. A. (2011) Identifying government spending shocks: It's all in the timing, *The Quarterly Journal of Economics*, **126**, 1–50.

Ramey, V. A. and Shapiro, M. D. (1998) Costly capital reallocation and the effects of government spending, *Carnegie-Rochester Conference Series on Public Policy*, **48**, 145–194.

Rotemberg, J. J. and Woodford, M. (1992) Oligopolistic pricing and the effects of aggregate demand on economic activity, *Journal of Political Economy*, **100**, 1153–1207.

Sims, C. A. (1998) Stickiness, *Carnegie-Rochester Conference Series on Public Policy*, **49**, 317–356.

Sims, C. A. (2003) Implications of rational inattention, *Journal of Monetary Economics*, **50**, 665–690.

Sims, C. A. (2005) Rational inattention: a research agenda, Tech. rep.

Sims, C. A. (2010) Rational inattention and monetary economics, in *Handbook of Monetary Economics* (Eds.) B. M. Friedman and M. Woodford, Elsevier, vol. 3 of *Handbook of Monetary Economics*, chap. 4, pp. 155–181.

Woodford, M. (2008) Information-constrained state-dependent pricing, NBER Working Papers 14620, National Bureau of Economic Research, Inc.

Technical Appendix

A. Derivation of the household's optimization problem

The derivation of the household's problem is almost identical to the derivation in the Appendix to Mackowiak and Wiederholt (2010). Nevertheless we present the solution in detail for the sake of consistency of this paper.

Under perfect information the household solves the problem 2.1 subject to 2.2 - 2.4. We guess (the guess will be verified) that the demand function for labor $d_w(W_{jt}, W_t, L_t)$ takes the form

$$L_{jt} = \left(\frac{W_{jt}}{W_t}\right)^{-\eta} L_t \tag{.1}$$

We define new variables

$$\hat{C}_{ijt} = \frac{C_{ijt}}{C_{jt}} \tag{.2}$$

Defining $\Gamma_t = 1 - \tau_t^w$ with $\tau_t = \log \Gamma_t - \log(1 - \tau^w)$ where τ^w is the steady state wage tax rate, we can express the budget constraint as

$$C_{jt} \sum_{i=1}^{I} P_{it}\hat{C}_{ijt} + B_{jt} = R_{t-1}B_{jt-1} + \frac{D_t}{J} - \frac{T_t}{J} + \Gamma_t W_{jt} L_{jt} \tag{.3}$$

We determine C_{jt} using the above equation

$$C_{jt} = \frac{R_{t-1}B_{jt-1} - B_{jt} + \frac{D_t}{J} - \frac{T_t}{J} + \Gamma_t W_{jt} L_{jt}}{\sum_{i=1}^{I} P_{it}\hat{C}_{ijt}} \tag{.4}$$

We divide the numerator and denominator by the aggregate price level P_t (the tilde denotes real variables, e.g. $\tilde{B}_{jt} = \frac{B_{jt}}{P_t}$)

$$C_{jt} = \frac{\frac{R_{t-1}}{\Pi_t}\tilde{B}_{jt-1} - \tilde{B}_{jt} + \frac{\tilde{D}_t}{J} - \frac{\tilde{T}_t}{J} + \Gamma_t \tilde{W}_{jt} L_{jt}}{\sum_{i=1}^{I} \hat{P}_{it}\hat{C}_{ijt}} \tag{.5}$$

Note that

$$1 = \sum_{i=1}^{I} \hat{C}_{ijt}^{\frac{\theta-1}{\theta}} \tag{.6}$$

Using all the above relations we can write the instantaneous utility function as

$$\begin{aligned}&\frac{1}{1-\gamma}\left(\frac{\frac{R_{t-1}}{\Pi_t}\tilde{B}_{jt-1} - \tilde{B}_{jt} + \frac{\tilde{D}_t}{J} - \frac{\tilde{T}_t}{J} + \Gamma_t \tilde{W}_{jt}\left(\frac{\tilde{W}_{jt}}{\tilde{W}_t}\right)^{-\eta} L_t}{\sum_{i=1}^{I-1} \hat{P}_{it}\hat{C}_{ijt} + \hat{P}_{It}\left(1 - \sum_{i=1}^{I-1} \hat{C}_{ijt}^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}}\right)^{1-\gamma} \\ &-\frac{1}{1-\gamma} - \frac{\varphi}{1+\psi}\left[\left(\frac{\tilde{W}_{jt}}{\tilde{W}_t}\right)^{-\eta} L_t\right]^{1+\psi}\end{aligned} \tag{.7}$$

Now we log-deviate the utility function

$$\begin{aligned}&\frac{1}{1-\gamma}\left(\frac{\frac{R\tilde{B}_j}{\Pi}e^{r_{t-1}-\pi_t+\tilde{b}_{jt-1}} - \tilde{B}_j e^{\tilde{b}_{jt}} + \frac{\tilde{D}}{J}e^{\tilde{d}_t} - \frac{\tilde{T}_t}{J}e^{\tilde{t}_t} + \Gamma\frac{\tilde{W}_j^{1-\eta}}{\tilde{W}^{-\eta}}Le^{(1-\eta)\tilde{w}_{jt}+\eta\tilde{w}_t+l_t+\tau_t}}{\frac{1}{I}\sum_{i=1}^{I-1} e^{\hat{p}_{it}+\hat{c}_{ijt}} + \frac{1}{I}e^{\hat{p}_{It}}\left(I - \sum_{i=1}^{I-1} e^{\frac{\theta-1}{\theta}\hat{c}_{ijt}}\right)^{\frac{\theta}{\theta-1}}}\right)^{1-\gamma} \\ &-\frac{1}{1-\gamma} - \frac{\varphi}{1+\psi}\left(\frac{\tilde{W}_j^{-\eta}}{\tilde{W}^{-\eta}}L\right)^{1+\psi} e^{-\eta(1+\psi)(\tilde{w}_{jt}-\tilde{w}_t)+(1+\psi)l_t}\end{aligned} \tag{.8}$$

We use steady state relations to get

$$\frac{C_j^{1-\gamma}}{1-\gamma}\left(\frac{\frac{\omega_B}{\beta}e^{r_{t-1}-\pi_t+\tilde{b}_{jt-1}}-\omega_B e^{\tilde{b}_{jt}}+\omega_D e^{\tilde{d}_t}-\omega_T e^{\tilde{t}_t}+\Gamma\omega_W e^{(1-\eta)\tilde{w}_{jt}+\eta\tilde{w}_t+l_t+\tau_t}}{\frac{1}{I}\sum_{i=1}^{I-1}e^{\hat{p}_{it}+\hat{c}_{ijt}}+\frac{1}{I}e^{\hat{p}_{It}}\left(I-\sum_{i=1}^{I-1}e^{\frac{\theta-1}{\theta}\hat{c}_{ijt}}\right)^{\frac{\theta}{\theta-1}}}\right)^{1-\gamma} \\ -\frac{1}{1-\gamma}-\frac{C_j^{1-\gamma}}{1+\psi}\frac{\eta-1}{\eta}\Gamma\omega_W e^{-\eta(1+\psi)(\tilde{w}_{jt}-\tilde{w}_t)+(1+\psi)l_t} \tag{.9}$$

where

$$(\omega_B,\omega_D,\omega_T,\omega_W)=(\frac{\tilde{B}_j}{C_j},\frac{\frac{\tilde{D}}{J}}{C_j},\frac{\frac{\tilde{T}}{J}}{C_j},\frac{L_j\tilde{W}_j}{C_j}) \tag{.10}$$

Next we derive a log-quadratic approximation of the objective function around the steady state. Let

$$x_t'=(\tilde{b}_{jt},\tilde{w}_{jt},\hat{c}_{1jt},...,\hat{c}_{I-1jt})$$

$$z_t'=(r_{t-1},\pi_t,\tilde{w}_t,l_t,\tilde{d}_t,\tilde{t}_t,\tau_t,\hat{p}_{1t},...,\hat{p}_{It})$$

x_t gathers choice variables and z_t gathers exogenous variables.

To include $\tilde{b}_{jt-1}$ let us define for convenience

$$x_{-1}'=(\tilde{b}_{j,-1},0,...,0)$$

Let g denote a function that we get by multiplying the instantaneous utility function by β^t for each $t\geq 0$ and summing up the obtained expressions from 0 to ∞ (discounted welfare). Let $\tilde{g}$ denote its second order approximation around the steady state. Then

$$E_{j,-1}\left[\tilde{g}(x_{-1},x_0,z_0,x_1,z_1,...\right]= \\ E_{j,-1}[g(0,0,...,0)+\sum_{t=0}^{\infty}\beta^t[h_x'x_t+h_z'z_t+\frac{1}{2}x_t'H_{x,-1}x_{t-1}+\frac{1}{2}x_t'H_{x,0}x_t+\frac{1}{2}x_t'H_{x,1}x_{t+1} \\ +\frac{1}{2}x_t'H_{xz,0}z_t+\frac{1}{2}x_t'H_{xz,1}z_{t+1}+\frac{1}{2}z_t'H_{z,0}z_t+\frac{1}{2}z_t'H_{zx,-1}x_{t-1}+\frac{1}{2}z_t'H_{zx,0}x_t] \\ +\beta^{-1}(h_{-1}'x_{-1}+\frac{1}{2}x_{-1}'H_{-1}x_{-1}+\frac{1}{2}x_{-1}'H_{x,1}x_0+\frac{1}{2}x_{-1}'H_{xz,1}z_0)] \tag{.11}$$

where the matrices H's and vectors $h's$ denote the appropriate derivatives. Under some technical conditions[26] (See the Appendix to Mackowiak and Wiederholt, 2010 for details) and using the fact that some of the derivatives are identically zero we can write

$$E_{j,-1}\left[\tilde{g}(x_{-1},x_0,z_0,x_1,z_1,...\right]=g(0,0,...,0)+ \\ \sum_{t=0}^{\infty}\beta^tE_{j,-1}[h_x'x_t]+\sum_{t=0}^{\infty}\beta^tE_{j,-1}[h_z'z_t]+\sum_{t=0}^{\infty}\beta^tE_{j,-1}[x_t'H_{x,1}x_{t+1}]+\sum_{t=0}^{\infty}\beta^tE_{j,-1}[\frac{1}{2}x_t'H_{x,0}x_t] \\ +\sum_{t=0}^{\infty}\beta^tE_{j,-1}[x_t'H_{xz,0}z_t]+\sum_{t=0}^{\infty}\beta^tE_{j,-1}[x_t'H_{xz,1}z_{t+1}]+\sum_{t=0}^{\infty}\beta^tE_{j,-1}[\frac{1}{2}z_t'H_{z,0}z_t] \\ +\beta^{-1}E_{j,-1}(h_{-1}'x_{-1}+\frac{1}{2}x_{-1}'H_{-1}x_{-1}+x_{-1}'H_{x,1}x_0+x_{-1}'H_{xz,1}z_0)] \tag{.12}$$

Now define the process $\{x_t^\star\}_{t=-1}^{\infty}$ which in every period $t\geq 0$ satisfies the first-order condition of maximizing .12 (E_t denotes the expectation operator conditioned on information available in period t)

$$E_t[h_x+H_{x,-1}x_{t-1}^\star+H_{x,0}x_t^\star+H_{x,1}x_{t+1}^\star+H_{xz,0}z_t+H_{xz,1}z_{t+1}]=0 \tag{.13}$$

[26]The conditions basically say that the variables in question cannot rise too fast in time

with $x^\star_{-1}$ given by $(\tilde{b}_{j,-1}, 0, ..., 0)$

Multiply .13 by $(x_t - x^\star_t)$ (this term is not random given information at time t)

$$E_t[(x_t - x^\star_t)'(h_x + H_{x,-1}x^\star_{t-1} + H_{x,0}x^\star_t + H_{x,1}x^\star_{t+1} + H_{xz,0}z_t + H_{xz,1}z_{t+1})] = 0$$

Using the law of iterated expectations we get

$$E_{j,-1}[(x_t - x^\star_t)'(h_x + H_{x,-1}x^\star_{t-1} + H_{x,0}x^\star_t + H_{x,1}x^\star_{t+1} + H_{xz,0}z_t + H_{xz,1}z_{t+1})] = 0$$

or

$$E_{j,-1}[(x_t - x^\star_t)'(h_x + H_{xz,0}z_t + H_{xz,1}z_{t+1})] = E_{j,-1}[(x_t - x^\star_t)'(H_{x,-1}x^\star_{t-1} + H_{x,0}x^\star_t + H_{x,1}x^\star_{t+1})] \quad (.14)$$

Now we will express the optimization problem of the household as minimizing the expected loss arising from a suboptimal choice x_t.

$$E_{j,-1}\left[\tilde{g}(x_{-1}, x_0, z_0, x_1, z_1, ...\right] - E_{j,-1}\left[\tilde{g}(x^\star_{-1}, x^\star_0, z_0, x^\star_1, z_1, ...\right] =$$

$$\sum_{t=0}^{\infty} \beta^t E_{j,-1}[x'_t H_{x,1}x_{t+1} + \frac{1}{2}x'_t H_{x,0}x_t - x^{\star\prime}_t H_{x,1}x^\star_{t+1} - \frac{1}{2}x^{\star\prime}_t H_{x,0}x^\star_t] +$$

$$\sum_{t=0}^{\infty} \beta^t E_{j,-1}[(x_t - x^\star_t)(h_x + H_{xz,0}z_t + H_{xz,1}z_{t+1})] +$$

$$\beta^{-1}E_{j,-1}(h'_{-1}x_{-1} + \frac{1}{2}x'_{-1}H_{-1}x_{-1} + \frac{1}{2}x'_{-1}H_{x,1}x_0 + \frac{1}{2}x'_{-1}H_{xz,1}z_0)] -$$

$$\beta^{-1}E_{j,-1}(h'_{-1}x^\star_{-1} + \frac{1}{2}x^{\star\prime}_{-1}H_{-1}x^\star_{-1} + x^{\star\prime}_{-1}H_{x,1}x^\star_0 + x^{\star\prime}_{-1}H_{xz,1}z_0)]$$

Using the expression .13 and the fact $x_{-1} = x^\star_{-1}$ we have

$$E_{j,-1}\left[\tilde{g}(x_{-1}, x_0, z_0, x_1, z_1, ...\right] - E_{j,-1}\left[\tilde{g}(x^\star_{-1}, x^\star_0, z_0, x^\star_1, z_1, ...\right] =$$

$$\sum_{t=0}^{\infty} \beta^t E_{j,-1}[x'_t H_{x,1}x_{t+1} + \frac{1}{2}x'_t H_{x,0}x_t - x^{\star\prime}_t H_{x,1}x^\star_{t+1} - \frac{1}{2}x^{\star\prime}_t H_{x,0}x^\star_t] -$$

$$\sum_{t=0}^{\infty} \beta^t E_{j,-1}[(x_t - x^\star_t)(H_{x,-1}x^\star_{t-1} + H_{x,0}x^\star_t + H_{x,1}x^\star_{t+1})] +$$

$$\beta^{-1}E_{j,-1}(x'_{-1}H_{x,1}(x_0 - x^\star_0))$$

Using equation .14 we finally arrive at

$$E_{j,-1}\left[\tilde{g}(x_{-1}, x_0, z_0, x_1, z_1, ...\right] - E_{j,-1}\left[\tilde{g}(x^\star_{-1}, x^\star_0, z_0, x^\star_1, z_1, ...\right] =$$

$$\sum_{t=0}^{\infty} \beta^t E_{j,-1}[\frac{1}{2}(x_t - x^\star_t)'H_{x,0}(x_t - x^\star_t) + (x_t - x^\star_t)'H_{x,1}(x_{t+1} - x^\star_{t+1})] \quad (.15)$$

After some tedious arithmetics required to compute the derivatives we are ready to characterize the solution to the houshold's problem under perfect information.

The loglinearized equation for c_{jt} (from 2.2) is

$$c_{jt} = \frac{\omega_B}{\beta}(r_{t-1} - \pi_t + \tilde{b}_{jt-1}) - \omega_B\tilde{b}_{jt} + \omega_D\tilde{d}_t - \omega_T\tilde{t}_t + \Gamma\omega_W((1-\eta)\tilde{w}_{jt} + \eta\tilde{w}_t + l_t + \tau_t)$$

Using this result and equation .13 we get the first equation characterizing the solution

$$c^\star_{jt} - E_{j,t}c^\star_{jt+1} = -\frac{1}{\gamma}\left(r_t - E_{j,t}\pi_{t+1}\right) \quad (.16)$$

The second equation is

$$\gamma c^{\star}_{jt} - (1+\eta\psi)\,\tilde{w}^{\star}_{jt} + \eta\psi\tilde{w}_t + \psi l_t - \tau_t = 0$$

or

$$\tilde{w}^{\star}_{jt} = \frac{\eta\psi}{1+\eta\psi}\tilde{w}_t + \frac{\psi}{1+\eta\psi}l_t + \frac{\gamma}{1+\eta\psi}c^{\star}_{jt} - \frac{1}{1+\eta\psi}\tau_t \tag{.17}$$

The rest $I-1$ equations are of the form

$$\hat{c}^{\star}_{ijt} + \sum_{k=1}^{I-1}\hat{c}^{\star}_{kjt} = -\theta(\hat{p}_{it} - \hat{p}_{It}) \text{ for } i = 1,2,...,I-1$$

Simplifying

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

After substitution

$$\hat{c}^{\star}_{ijt} = -\theta\left(\hat{p}_{it} - \frac{1}{I}\sum_{i=1}^{I}\hat{p}_{it}\right) \tag{.18}$$

Note that equation .18 enables us to derive the demand function for different consumption goods. 2.22 and 2.24 imply that $\sum_{i=1}^{I}\hat{p}_{it} = 0$. Hence

$$\hat{c}^{\star}_{ijt} = c^{\star}_{ijt} - c^{\star}_{jt} = -\theta\hat{p}_{it} \tag{.19}$$

Summing over j and dividing by J yields[27]

$$c_{it} - c_t = -\theta\hat{p}_{it} \tag{.20}$$

Coming back to original variables

$$C_{it} = \hat{P}_{it}^{-\theta}C_t = d_c(P_{it}, P_t, C_t) \tag{.21}$$

We can now derive the optimization problem of the wage-setter 2.42 - 2.46 and consumer 2.47 - 2.54. The objective function is given by .15. Note that the way we derived .15 implies that this expression is always negative and equals zero if and only if the agent chooses the optimal bond holding, wage rate and consumption mix. Since the expression $E_{j,-1}\left[\tilde{g}(x^{\star}_{-1}, x^{\star}_0, z_0, x^{\star}_1, z_1, ...\right]$ is exogenous from the point of view of the household, maximizing .15 is equivalent to maximizing $E_{j,-1}\left[\tilde{g}(x_{-1}, x_0, z_0, x_1, z_1, ...\right]$. Instead of treating the total level of attention, κ, as a fixed parameter, we let the household choose it optimally. Therefore we subtract the welfare cost of information $\frac{\lambda}{1-\beta}\kappa$ from .15[28] arriving at 2.42 and 2.47. The only difference between these two expressions is that in the first case we let the agent choose the wage rate and treat other choice variables as given and in the second case we treat the wage rate as given and let the agent choose the remaining variables.

We get equation 2.44 from .17, equation 2.50 from .16 and equation 2.52 from .19. The only condition is that we can express those variables as a MA process based on the innovations ε^A_t, ε^G_t driving the aggregate shocks. This in turn is possible if such a representation can be found for aggregate variables r_t, π_t, $\tilde{w}_t$, l_t and τ_t. We show that this is possible later when we characterize the general equilibrium of the model under perfect information. Now we restrict the choice of wage rate and composite consumption to the form 2.45 and 2.51 and the choice of consumption mix to the form 2.53. This means that the actual choice can be seen as a sum of a pure signal and noise. The existence of noise allows the agent to regulate the precision and informational content of his actual decision which in turn determines the value of the objective function and the corresponding information flow. Note that we exchanged the choice of bond holdings for the choice of composite consumption. This is

[27] Note that after loglinearization $c_{it} = \frac{1}{J}\sum_{j=1}^{J}c_{ijt}$

[28] One has to admit that this way of including the cost of information in the objective function is a little bit ad hoc. However, it has a big advantage that utility is linear in the attention level κ, which greatly simplifies solving the model.

allowed thanks to equation 2.49 which follows directly from the budget constraint and the assumption that initial bond holding is zero.

A natural question arises whether restricting the decision to the form 2.45 and 2.51 is binding, or put differently, whether such form (in particular the Gaussianity of noise) is optimal. Unfortunately we don't know the answer to that question. Some insight is provided by a result proven by Mackowiak and Wiederholt **?** that the form: pure signal plus Gaussian noise is optimal when the objective function is quadratic and the optimal choice follows an AR(1) process.

Finally, we add the information processing constraint 2.46 and 2.54 completing the specification of the maximization problem of the household. We use mutual information 2.33 to quantify information flow between the actual and optimal decision. Since we assumed that choosing the wage rate and choosing composite consumption and consumption mix are separate activities, we can divide the information flow constraint into two independent constraints: one restricting the information flow in the wage-setter problem and one restricting information flow in the consumer problem.

B. Derivation of the firm's optimization problem

Since the derivation of the firm's problem is analogous to the derivation of the household's problem we will skip most of the details and only sketch the main results.

The firm solves the following problem

$$\max_{\{L_{i1t},...,L_{iJt},P_{it}\}_{t=0}^{\infty}} E_{i,-1}\left\{\sum_{t=0}^{\infty} Q_t\left(P_{it}Y_{it}-\sum_{j=1}^{J} W_{jt}L_{ijt}\right)\right\} \tag{.22}$$

subject to

$$Y_{it}=e^{a_t}e^{a_{it}}L_{it}^{\alpha} \tag{.23}$$

$$Y_{it}=G_{it}+C_{it} \tag{.24}$$

$$C_{it}=d_c(P_{it},P_t,C_t) \tag{.25}$$

$$G_{it}=d_g(P_{it},P_t,G_t) \tag{.26}$$

We use the assumption 2.19, 2.11 equation .21 and introduce variables $\hat{L}_{ijt}=\frac{L_{ijt}}{L_{it}}$. After some algebra we arrive at the unconstrained, equivalent problem

$$\max_{\{\hat{L}_{i1t},...,\hat{L}_{i(J-1)t},\hat{P}_{it}\}_{t=0}^{\infty}} E_{i,-1}\sum_{t=0}^{\infty}\beta^t\Lambda(...)\left\{\hat{P}_{it}^{1-\theta}(C_t+G_t)-\left[\frac{\hat{P}_{it}^{-\theta}(C_t+G_t)}{e^{a_t}e^{a_{it}}}\right]^{\frac{1}{\alpha}}\left[\sum_{j=1}^{J-1}\tilde{W}_{jt}\hat{L}_{ijt}+\tilde{W}_{Jt}\left(1-\sum_{j=1}^{J-1}\hat{L}_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\right]\right\} \tag{.27}$$

where $\Lambda(...)=\Lambda(C_{1t},...,C_{Jt})$.

Next we log-deviate the objective function (denote this function f) and use its second order approximation ($\tilde{f}$). Let

$$x_t'=(\hat{p}_{it},\hat{l}_{i1t},...,\hat{l}_{i(J-1)t}) \tag{.28}$$

$$z_t'=(a_t,a_{it},c_t,g_t,\tilde{w}_{1t},..,\tilde{w}_{Jt}) \tag{.29}$$

Then

$$\tilde{f}(x_t,z_t)=f(0,0)+h_x'x_t+h_z'z_t+\frac{1}{2}x_t'H_xx_t+\frac{1}{2}z_t'H_zz_t+x_t'H_{xz}z_t \tag{.30}$$

Under conditions analogous to those presented in the derivation of the household's objective we have

$$E_{i,-1}\sum_{t=0}^{\infty}\beta^t\left(f(0,0)+h_x'x_t+h_z'z_t+\frac{1}{2}x_t'H_xx_t+\frac{1}{2}z_t'H_zz_t+x_t'H_{xz}z_t\right)=$$

$$=\sum_{t=0}^{\infty}\beta^tE_{i,-1}\left(f(0,0)+h_x'x_t+h_z'z_t+\frac{1}{2}x_t'H_xx_t+\frac{1}{2}z_t'H_zz_t+x_t'H_{xz}z_t\right)$$

The necessary and sufficient condition for optimal x_t, denoted $x_t^\star$, is

$$h_x' + H_x x_t^\star + H_{xz} z_t = 0 \tag{.31}$$

Using $x_t^\star$ we express the objective of the firm as minimizing the expected loss from suboptimal decision

$$E_{i,-1}\left[\sum_{t=0}^{\infty}\beta^t f(x_t, z_t) - \sum_{t=0}^{\infty}\beta^t f(x_t^\star, z_t)\right] = \sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}(x_t - x_t^\star)' H_x (x_t - x_t^\star)\right] \tag{.32}$$

Computing the derivatives and using steady state relations enables us to characterize optimal solution according to .31. The first equation of the system .31 reads

$$\hat{p}_{it}^\star = -\frac{\frac{1}{\alpha}}{1+\theta\frac{1-\alpha}{\alpha}}(a_t + a_{it}) + \frac{\theta_c\frac{1-\alpha}{\alpha}}{1+\theta\frac{1-\alpha}{\alpha}}c_t + \frac{\theta_g\frac{1-\alpha}{\alpha}}{1+\theta\frac{1-\alpha}{\alpha}}g_t + \frac{1}{1+\theta\frac{1-\alpha}{\alpha}}\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt} \tag{.33}$$

The remaining $J-1$ equations are of the form

$$\hat{l}_{ijt}^\star + \sum_{k=1}^{J-1}\hat{l}_{ikt}^\star = -\eta(\tilde{w}_{jt} - \tilde{w}_{Jt}) \text{ for } j = 1, 2, ..., J-1 \tag{.34}$$

Simplifying

$$\hat{l}_{ijt}^\star = -\eta\hat{w}_{jt} \tag{.35}$$

Using the facts $\frac{1}{J}\sum_{j=1}^{J}\tilde{w}_{jt} = \tilde{w}_t$, $\hat{w}_{jt} = \tilde{w}_{jt} - \tilde{w}_t$, summing over i and dividing by I yields

$$l_{jt} - l_t = -\eta\hat{w}_{jt} \tag{.36}$$

or in the original variables

$$L_{jt} = \hat{W}_{jt}^{-\eta} L_t = d_w(W_{jt}, W_t, L_t) \tag{.37}$$

Thereby we have verified the guess .1.

Collecting .32, .33, .35, introducing the information flow constraint, subtracting the cost of information $\frac{\mu}{1-\beta}\kappa$ from the firm's objective and restricting the actual decisions to 2.39 and 2.40 we arrive at the problem 2.35 - 2.41.

C. The equilibrium of the model with perfect information

In this section we characterize the equilibrium dynamics of an economy in which all agents have perfect information.

We start by deriving the equations describing the behavior of fiscal and monetary authority. Dividing 2.13 on both sides by P_t, using the definition $\Gamma_t = 1 - \tau_t^w$ and $\tau_t = \log\Gamma_t - \log(1-\tau^w)$ and loglinearizing yields

$$\frac{\theta_b}{\beta}(r_{t-1} - \pi_t + \tilde{b}_{t-1}) + \theta_g g_t = \theta_b\tilde{b}_t + \tau^w\theta_{WL}(\tilde{w}_t + l_t) - (1-\tau^w)\theta_{WL}\tau_t + \theta_T\tilde{t}_t$$

where

$$\theta_{WL} = \frac{\tilde{W}L}{Y},\ \theta_b = \frac{\tilde{B}}{Y},\ \theta_T = \frac{\tilde{T}}{Y}$$

θ_T can be determined using the fixed parameter θ_τ

$$\theta_T = \tau^w\theta_{WL}\frac{1-\theta_\tau}{\theta_\tau}$$

θ_b is related to ω_B in the following way

$$\omega_B = \frac{\tilde{B}_j}{C_j} = \frac{\tilde{B}}{C} = \frac{\theta_b}{\theta_c}$$

θ_{WL} can be calculated directly from the steady state values of $\tilde{W}$, L and Y which we don't report here.

Loglinearizing 2.14 yields

$$S^{\star}\theta_g(s_t + g_t) = \tau^w\theta_{WL}(\tilde{w}_t + l_t) - (1-\tau^w)\theta_{WL}\tau_t + \theta_T\tilde{t}_t$$

where $S^{\star}$ is determined endogenously from

$$S^{\star} = \frac{\tau^w\theta_{WL} + \theta_T}{\theta_g}$$

To get rid of $\tilde{t}_t$ we use the assumption

$$\frac{\theta_\tau}{1-\theta_\tau} = \frac{\tau_t^w\tilde{W}_t L_t}{\tilde{T}_t}$$

which implies

$$\theta_T\tilde{t}_t = \frac{1-\theta_\tau}{\theta_\tau}\left(\tau^w\theta_{WL}(\tilde{w}_t + l_t) - (1-\tau^w)\theta_{WL}\tau_t\right)$$

As a result we get

$$S^{\star}\theta_g(s_t + g_t) = \tau^w\frac{\theta_{WL}}{\theta_\tau}(\tilde{w}_t + l_t) - (1-\tau^w)\frac{\theta_{WL}}{\theta_\tau}\tau_t$$

and the budget constraint

$$\tilde{b}_t = \frac{1}{\beta}(r_{t-1} - \pi_t + \tilde{b}_{t-1}) + \frac{\theta_g}{\theta_b}g_t - \frac{\tau^w\theta_{WL}}{\theta_\tau\theta_b}(\tilde{w}_t + l_t) + \frac{(1-\tau^w)\theta_{WL}}{\theta_\tau\theta_b}\tau_t$$

Loglinearizing 2.15 we get

$$S^{\star}s_t = \omega y_t$$

As a last step we express the tax rate as a function of g_t, c_t, l_t and $\tilde{w}_t$

$$\tau_t = -\frac{(\omega\theta_g + S^{\star})\,\theta_\tau\theta_g}{(1-\tau^w)\theta_{WL}}g_t - \frac{\omega\theta_\tau\theta_g\theta_c}{(1-\tau^w)\theta_{WL}}c_t + \frac{\tau^w}{1-\tau^w}(\tilde{w}_t + l_t)$$

Loglinearizing the monetary policy rule 2.21 yields

$$r_t = \rho_R r_{t-1} + (1-\rho_R)(\phi_\pi\pi_t + \phi_y y_t) \tag{.38}$$

Now we will use the results of households' and firms' optimization to complete the description of the equilibrium.

Equation .33 together with $\sum_{i=1}^{I}\hat{p}_{it} = 0$, $w_t = \frac{1}{J}\sum_{j=1}^{J}w_{jt}$ and $\sum_{i=1}^{I}a_{it} = 0$ implies

$$0 = \theta_c(1-\alpha)c_t + \theta_g(1-\alpha)g_t + \alpha\tilde{w}_t - a_t \tag{.39}$$

Using .17, summing across j and dividing by J yields[29]

$$\tilde{w}_t = \psi l_t + \gamma c_t - \tau_t \tag{.40}$$

Equating production 2.6 with demand 2.8 and aggregating allows us to express labor supply as

$$l_t = \frac{\theta_c}{\alpha}c_t + \frac{\theta_g}{\alpha}g_t - \frac{1}{\alpha}a_t \tag{.41}$$

[29] When interpreting this equation one should remember that the definition $\Gamma_t = 1 - \tau_t^w$ implies that τ_t falls when the tax rate τ_t^w rises.

The above equations are sufficient to express consumption as a function of exogenous variables and τ_t

$$c_t = \frac{1+\psi}{\theta_c(1-\alpha+\psi)+\alpha\gamma}a_t - \frac{\theta_g(1-\alpha+\psi)}{\theta_c(1-\alpha+\psi)+\alpha\gamma}g_t + \frac{\alpha}{\theta_c(1-\alpha+\psi)+\alpha\gamma}\tau_t \tag{.42}$$

Aggregating the Euler equation .16 we get

$$c_t - E_t c_{t+1} = -\frac{1}{\gamma}\left(r_t - E_t\pi_{t+1}\right) \tag{.43}$$

After some more algebra we arrive at a system of equations that fully characterizes the dynamics of the economy

$$c_t - E_t c_{t+1} = -\frac{1}{\gamma}\left(r_t - E_t\pi_{t+1}\right) \tag{.44}$$

$$r_t = \rho_R r_{t-1} + (1-\rho_R)(\phi_\pi \pi_t + \phi_y \theta_c c_t + \phi_y \theta_g g_t) \tag{.45}$$

$$a_t = \rho_A a_{t-1} + \varepsilon_t^A \tag{.46}$$

$$g_t = \rho_G g_{t-1} + \varepsilon_t^G \tag{.47}$$

$$c_t = \Psi_1 a_t - \Psi_2 g_t \tag{.48}$$

where

$$\Psi_1 = \frac{\theta_g(1-\tau^w)(1+\psi) - \alpha\theta_g\left(1-\frac{(\omega\theta_g+S^\star)\theta_\tau}{\theta_{WL}}\right)}{(1-\tau^w)\left(\theta_c(1+\psi)+\alpha\gamma\right) - \alpha\theta_c\left(1-\omega\frac{\theta_g\theta_\tau}{\theta_{WL}}\right)}$$

$$\Psi_2 = \frac{(1-\tau^w)(1+\psi)}{(1-\tau^w)\left(\theta_c(1+\psi)+\alpha\gamma\right) - \alpha\theta_c\left(1-\omega\frac{\theta_g\theta_\tau}{\theta_{WL}}\right)}$$

D. Optimal consumption and factor mix under rational inattention

In this section we solve analytically a part of the consumer's optimization problem, namely the choice of the optimal consumption mix 2.53. This subproblem can be singled out from the problem 2.47 - 2.54, because of the following facts: (i) relative prices $\hat{p}_{it}$ and noise innovations ξ_{ijt}^C are independent from all aggregate variables and other shocks, (ii) mutual information is additive in independent variables, (iii) the objective function is linear in attention capacity κ, (iv) matrices H_0 and H_1 have a block structure with zeros in the top right and bottom left corners (this is easy to verify). The last condition says that the decision regarding the consumption mix does not influence the decision regarding bond holding (or composite consumption) and vice versa.

Using steady state relations and the elements of matrices H_0 and H_1 we can write the problem as

$$\max_{\tilde{\theta},\,\vartheta}\left\{-C_j^{1-\gamma}\frac{1}{\theta I}\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\sum_{1\le k\le l\le I-1}(\hat{c}_{kjt}-\hat{c}_{kjt}^\star)(\hat{c}_{ljt}-\hat{c}_{ljt}^\star)\right] - \frac{\lambda}{1-\beta}\kappa_c\right\}$$

subject to

$$\hat{c}_{ijt}^\star = -\theta\hat{p}_{it} \quad i = 1,2,...,I-1 \tag{.49}$$

$$\hat{c}_{ijt} = -\tilde{\theta}\left(\hat{p}_{it} + \vartheta\xi_{ijt}^C\right) \quad i = 1,2,...,I-1 \tag{.50}$$

$$\mathcal{I}\left(\left\{\hat{c}_{1jt}^\star,...,\hat{c}_{(I-1)jt}^\star\right\};\left\{\hat{c}_{1jt},...,\hat{c}_{(I-1)jt}\right\}\right) \le \kappa_c$$

Using independence[30] and symmetry and denoting $\sigma^2 = Var(\hat{p}_{it})$ we can simplify the problem to

$$\max_{\tilde{\theta},\,\vartheta}\left\{-C_j^{1-\gamma}\frac{1}{\theta}\frac{I-1}{I}\left((\theta-\tilde{\theta})^2\sigma^2 + \tilde{\theta}^2\vartheta^2\right) - \lambda\kappa_c\right\} \tag{.51}$$

[30] Note that equations .33 and .39 imply that $\hat{p}_{it}^\star = -\frac{\frac{1}{\alpha}}{1+\theta\frac{1-\alpha}{\alpha}}a_{it}$ and a_{it}'s are assumed independent for different $i's$. However the actual $\hat{p}_{it}$ under rational inattention may be different than $\hat{p}_{it}^\star$. We therefore assume that $\hat{p}_{it}$ are independent. This simplifying assumption keeps the problem tractable but is criticized by Sims (2010).

subject to

$$(I-1)\,\mathcal{I}\left(\left\{\tilde{\theta}\left(\hat{p}_{it}+\vartheta\xi_{ijt}^{C}\right)\right\};\{\theta\hat{p}_{it}\}\right)\leq\kappa_{c} \tag{.52}$$

In this simple case we can derive information flow (using 2.34) explicitly

$$\mathcal{I}\left(\left\{\tilde{\theta}\left(\hat{p}_{it}+\vartheta\xi_{ijt}^{C}\right)\right\};\{\theta\hat{p}_{it}\}\right)=\frac{1}{2}\log_{2}\left[\frac{\vartheta^{2}+\sigma^{2}}{\vartheta^{2}}\right] \tag{.53}$$

This result allows us to write the problem as a simple unconstrained maximization

$$\max_{\tilde{\theta},\,\vartheta}\left\{-C_{j}^{1-\gamma}\frac{1}{\theta}\frac{I-1}{I}\left((\theta-\tilde{\theta})^{2}\sigma^{2}+\tilde{\theta}^{2}\vartheta^{2}\right)-\frac{\lambda}{2}\log_{2}\left[\frac{\vartheta^{2}+\sigma^{2}}{\vartheta^{2}}\right]\right\}$$

It turns out after some algebra that the solution is

$$\tilde{\theta}=\theta\frac{\sigma^{2}}{\sigma^{2}+\vartheta^{2}}$$

$$\vartheta^{2}=\frac{\frac{\lambda}{2\ln 2}\sigma^{2}}{\left(C_{j}^{1-\gamma}\frac{\theta}{I}\sigma^{2}-\frac{\lambda}{2\ln 2}\right)}$$

provided that $C_{j}^{1-\gamma}\frac{\theta}{I}\sigma^{2}-\frac{\lambda}{2\ln 2}>0$. Otherwise (when the cost of information is sufficiently high relative to the variance of $\hat{p}_{it}$) it is optimal to choose $\vartheta^{2}=\infty$ which means no information flow.

Figure .1: Response to 1% shock to aggregate technology